

A Time Domain Based Classifier for ECG Pattern Recognition

G.P. Shorten, *Member IEEE* and M.J. Burke

Abstract—Pattern recognition, and in particular dynamic time warping has been applied to the ECG for many different purposes over the last decade. Significant research on creating adaptive, feature based, and more complex forms of the algorithm in order to increase its ability to classify an ECG signal accurately has been performed. Despite this increase in complexity and in the number of variations of the dynamic time warping algorithm there has been less focus on actually using the results of dynamic time warping to relate the reference and test signals to each other as accurately as possible. The majority of dynamic time warping algorithms published in the literature, even the most complex of them, classify the most accurate match to a reference signal based only on resulting Euclidean distance or slope difference between samples of the known reference and unknown query signal. This article demonstrates how a combination of measurements including heart-rate, amplitude and required warping time alignment can be used to reduce the resulting error in the classification of a query signal after the query and reference signals have been warped together. Its benefits are verified with significant testing.

I. INTRODUCTION

ECG pattern recognition has been at the forefront of cardiac research for a number of years. Although it has many applications, the fundamental objective in each application remains the same i.e. to identify characteristics within an ECG frame. One method of pattern recognition, originally emerging from the domain of speech analysis, is dynamic time warping (DTW) [1]. Because the method originates from an area of signal processing focusing on non-stationary signals obtained from non-linear systems it is no surprise that it has been applied to the ECG signal. The DTW process can take two signals of different frame length and duration and align them using a non-linear dynamic process. A single recursive measure of Euclidean distance or other difference measurements can be used to warp the signals in time so as to reduce the difference between them.

Different forms of dynamic time warping can be loosely grouped into three types; derivative, value and feature based dynamic time warping. Each has its own strengths and weaknesses associated with it and they have been combined to see the algorithm applied to ECG frame segmentation [2], classification [3] and even biometric identification [4]. The

DTW algorithm aligns the two signals in time to reduce the difference between the two signals be it on a value, derivative or feature-based premise.

The process of establishing which reference signal the test signal or query is most similar to is known as classification. There are a number of methods used to classify the query signal. Many authors [2] and [5] use a multilayered approach whereby when warping the signals the ECG component reference identified as the closest match is the one which provides the minimum slope differences between approximation splines during warping. Others such as [3] use the residual difference between the query and reference to perform the classification of the signals post-DTW. Although these methods of classification are popular, they have the limitation of being dependent on just one feature of the query and reference signals i.e. the amplitude or approximation spline slope differences between reference and query.

To overcome this limitation some users [4] have combined other analysis techniques with DTW such as Fisher's Linear Discriminant Analysis to further increase the accuracy of the classification process.

In this article it shall be shown how using a classifier that takes into account the query and reference signal heart rates, the amount of time alignment required around the fiducial points of the ECG signal and the resulting residual amplitude difference between the query and reference signals can aid in the classification process. The classification is not significantly complex, as all of the features used are in the time domain and hence, is not computationally intensive.

II. BACKGROUND

A. Dynamic Time Warping

The type of dynamic time warping used for this exercise is a value based algorithm, most similar to that used by Huang et al [3], although the classifier can be used with results from other methods of DTW. Value based DTW calculates the minimum Euclidean distance between each specific sample in the query and every sample in the reference signal.

Using the input signals two matrices of the same size are created, S_1 an $m \times n$ matrix which contains the reference signal repeated on each row and S_2 an $m \times n$ matrix which contains the query signal repeated in each column. A Euclidean distance matrix D can be calculated using single dimension Euclidean distance as given by equation (1):

$$D(x,y) = [S1(x,y) - S2(x,y)]^2 \quad (1)$$

where $1 \leq x \leq m$ and $1 \leq y \leq n$.

Manuscript received March, 2011. This work was supported in part by the Irish Research Council for Science, Engineering and Technology.

G.P. Shorten is a PhD research student with the Department of Electronic and Electrical Engineering, Trinity College Dublin, Dublin 2, Ireland (e-mail shorteg@tcd.ie).

M. J. Burke is a Senior Lecturer and Head of Department for the Department of Electronic and Electrical Engineering, Trinity College Dublin, Dublin 2, Ireland (e-mail martin.burke@tcd.ie).

Using the distance matrix D , a cumulative distance or C matrix can be created that measures the minimum cost of matching each sample in the reference and query signals. The cost matrix C is created by starting at location $(1,1)$ in the matrix D and calculating the cumulative distance of the first column and row of the matrix D storing the results in the corresponding location of the new matrix C (therefore an $m \times n$ matrix). The remaining cumulative values to be stored in the matrix C are calculated by following the recursive equation:

$$C(x,y) = d(x,y) + \min \begin{cases} d(x,y-1) \\ d(x-1,y-1) \\ d(x-1,y) \end{cases} \quad (2)$$

where $1 \leq x \leq m$ and $1 \leq y \leq n$.

In speech signal synthesis, where the DTW algorithm originated from it may not have been necessary to time align the two complete frames in their entirety. However, in ECG signal analysis one of the constraints usually introduced to the DTW process is ensuring the beginning and end of the frames are time aligned. This is ensured by starting at the last location $C(x,y)$ of the cumulative distance matrix C and moving to the lowest "cost" value stored in any one of the adjoining locations. By tracing all the way back to the beginning of the matrix C , and recording the path used that resulted in the minimum accumulated difference, two new time aligned sample sets are created called S_{1w} and S_{2w} . The x co-ordinates of the path are used to create S_{1w} and the y co-ordinates for S_{2w} . If a sample is repeated in order to time align the two signals it is referred to as padding.

B. Heart Rate of the Query and Reference Signals

The classifier consists of three time domain features with respect to the query and reference signals. Each must be expressed as a fraction before being used in the classifier equation. The first feature to be included is the relationship between the heart rate of the original signal and the reference signal. The difference between the heart rate of the current query and reference shall be expressed as a fraction of the maximum heart rate difference between the current query signal and the reference with the least similar heart rate from the test database. The heart rate variable X is given by:

$$X = \frac{|Q_{HR} - R_{HR}|}{\max |Q_{HR} - R_{dHR}|} \quad (3)$$

Where Q_{HR} and R_{HR} are the heart rates of the two signals under test, and R_{dHR} is the difference in beats per minute between the current query heart rate and least similar reference signal heart rate in the reference database.

C. Time Alignment Required around the Fiducial Points

The second feature to be included in the classification criteria accounts for the alteration of the test signals time scale around the fiducial points during time alignment. A fiducial point refers to the onset or termination point of one

of the constituent P, QRS or T-waves within an ECG signal. Since the duration of each constituent wave and hence the location of these points is considered diagnostically significant [8] and their identification is often the objective of applying DTW in the first place, we would like the query and reference signal to have been similar to each other in these regions before alignment. Of course it is possible that the morphologies are similar but require padding to time align them (as is the purpose of the DTW algorithm), but significant padding would imply that the morphologies of the two signals were in fact very different. This variable Y accounts for that possibility. It is calculated for each constituent wave of the current signals under test and is defined as the total padding inserted within ± 3 samples of the onset and termination of the P, QRS and T-waves of the current query and reference signal during DTW, expressed as a fraction of the maximum padding inserted during warping of the current query to the reference signal in the database which required the most time alignment around the fiducial points during DTW.

$$Y = \frac{|P_O + P_T|}{\max (|P_{RefO} + P_{RefT}|)} \quad (4)$$

Where P_O and P_T are the padding or number of samples inserted within ± 3 samples of the onset and termination of the current query and reference signal respectively. P_{RefO} and P_{RefT} are the total padding required during time alignment between the current query and the reference signal from the database that required the maximum amount of time alignment around the fiducial points.

D. RMS Difference of Query and Reference

The purpose of this feature in the classifier is similar to the residual or difference measurements used in [2] - [5]. The 'R' peaks of the query and reference signals should be amplitude normalized to the same value pre-DTW. After dynamic time warping has been performed one can measure the remaining amplitude difference between the query and reference signal and take that as a measure of similarity between the two signals. In this case we shall use a root-mean-square difference (RMS-difference) calculation to measure the amplitude difference of the time aligned query and reference signals. For each of the constituent waves of the signal i.e. the P, QRS and T-waves of the beat under analysis a RMS-difference variable Z will be calculated.

$$Z = \sqrt{\frac{\sum_{i=k}^n [R(i) - Q(i)]^2}{\sum_{i=k}^n R(i)^2}} \quad (5)$$

Where Q and R are the query and reference signals respectively, k is the onset sample number and n the termination sample number of the ECG P, QRS or T-wave being processed. The RMS-difference offers an insight into

the similarities of the two signals in amplitude terms after they have been time aligned.

E. The Time Domain Classifier Equation

The classifier uses a root of the sum of the squares for each of the features X, Y and Z to rank each of the results between the query and the signals in the reference database as shown in equation (6):

$$C_R = \sqrt{C_1(X^2) + C_2(Y^2) + C_3(Z^2)} \quad (6)$$

Where C_R is the resulting classifier rank, X, Y and Z are the feature variables and C_1 , C_2 and C_3 are scaling coefficients used to weight the contribution of each feature to the classifier. A classifier rank is created for the P, QRS and T wave of each query to reference comparison and the one which yields the minimum ranking is deemed the best match.

F. The Test Database

To examine the accuracy of the classifier and contribution of each feature a large database of test signals is required. The test data were taken from the MIT QT Database available for download online [6]. A total of 719 beats were chosen from the Lead II signals belonging to subsets of the Normal Sinus Rhythm and Arrhythmia databases. The original data from the QT database had a 250Hz sampling frequency. Annotations for the records were available for the fiducial points of each P, QRS and T-wave.

G. Measuring the Accuracy of the Process

In order to quantify the accuracy of the DTW process and the best reference classifier we shall use the location of the fiducial points as a metric. The duration of constituent P, QRS and T-waves are frequently used in cardio-diagnostics so the ability of the process to identify them in a query signal is of utmost importance. Each signal from the reference database shall be used as a query and warped to the other 718 reference signals. In doing so, we know the location of the fiducial points of the query as annotated by a cardiologist in the QT Database and can compare them with the fiducial points identified using the best match reference signal identified by the classifier after DTW.

III. OPTIMIZING THE SCALING COEFFICIENTS C_1 , C_2 AND C_3

In order to find the most accurate combination of scalars to weight the contribution of each feature, the coefficients C_1 , C_2 and C_3 shall be adjusted from 0 to 1 in 0.167 increments. Starting with C_3 , C_2 and finally C_1 all combinations of the scaling weight coefficients can be tested.

Fig. 1 is a plot of the total root mean square error (RMSE) in the identification of the fiducial points in all of the 719 test signals for each combination of the scaling coefficients.

It can be clearly seen that each feature has an effect on the accuracy of the classifier. Note how after the 49th combination, where C_1 (X) is increased from 0 to 0.167, there is a large reduction in the RMSE because the heart rate constraint has been introduced to the classifier. One can also

see that at every 49th combination from here on the error spikes to a peak as C_2 (Y), the measurement of padding around the fiducial point is removed when C_2 returns to zero. The final constraint Z, the RMS-difference or residual between the query and reference constituent P, QRS and T-waves is reset to zero every seventh combination as the coefficient C_3 is returned to zero. By observation of the resulting RMSE in Fig. 1, it was found that to achieve minimum RMSE (providing minimum error and maximum stability) a coefficient combination $C_1=0.333$, $C_2=0.833$ and $C_3=1.0$ should be used to scale or weight the contribution of each feature to the classifier.

This optimum combination of weight coefficients was discovered by test on the large reference database used in this exercise. If the process was repeated with a larger database with a larger variety of morphologies at different heart rates we might expect the weights to vary accordingly.

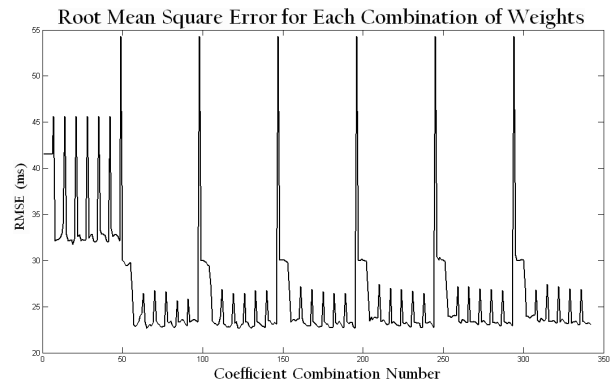


Fig. 1: Training the feature weight coefficients.

IV. MEASURING THE ACCURACY OF THE PROCESS

The test signals originating from the Normal Sinus Rhythm and Arrhythmia Databases are examined separately since they offer significantly differing morphologies. The algorithm is tested for its ability to locate fiducial points and resulting P, QRS and T-wave durations. Diagnoses of various cardiac conditions such as hypokalaemia [8] are based on the duration of constituent waves, signifying their importance. Note that 1.5% of the query/reference matches were deemed as obvious errors and removed before the subsequent analysis shown here.

A. Finding the Fiducial Points

The results in Table 1 are the mean and standard deviation of error between the actual location of the fiducial points and those identified with DTW. Also included are deviations regarded as acceptable by different cardiologist assessors of the same points as reported by Jané et al [7] which serve as a bench mark for performance of the automatic algorithm.

Table 1 demonstrates that the mean errors resulting from the algorithm are low, particularly if one considers that with a sampling frequency of 250Hz a mean error of 4ms is just one sample. Comparing the standard deviations of error for the algorithm with the deviation of expert cardiologist opinion one can see that the algorithm deviations are within $\pm 12\text{ms}$ or ± 3 samples. The algorithm deviations are higher

for some of the fiducial points. This is because the expert deviations are those of experts viewing of the same beats while in the case of the algorithm it is matching a different annotated reference beat.

Fiducial Point	Expert (ms)	Normal-Sinus Rhythm (ms)	Arrhythmia (ms)
P-Onset	±10.2	-2.05±17.23	-0.64±16.44
P-Termination	±12.7	2.48±13.30	5.23±17.47
QRS-Onset	±6.5	0.71±11.20	2.11±18.38
QRS-Termination	±11.6	1.75±10.83	5.07±14.78
T-Onset	N/A	-11.39±29.0	12.1±31.71
T-Termination	±30.6	6.76±14.87	6.03±27.62

Table 1: Mean ± standard deviation of the algorithm and classifier error compared with expected expert deviation.

B. Measuring the Duration of the Constituent Waves

To further investigate the accuracy of the DTW classifier process we investigate how well the duration of the P, QRS and T-waves from each signal can be measured.

In Fig. 2, the duration of each component as annotated by the expert cardiologists for each wave has been averaged over 6-bpm intervals. Also shown is the mean ± one standard deviation of the durations for each heart rate interval to give an idea of the variation in how the experts annotated the different waves for the same heart rate interval. We would hope the results of the warping process would lie within the deviations of the durations as annotated by experts.

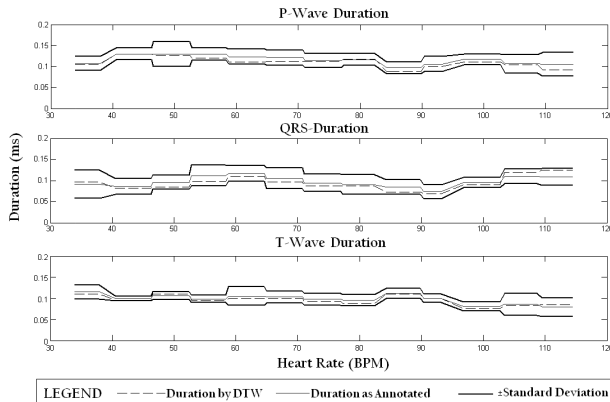


Fig.2: Arrhythmia Constituent Wave Durations as detected by DTW.

The results in Fig. 2 prove that the combination of value based DTW and our time domain classifier enable a very accurate automatic measurement of the duration of each

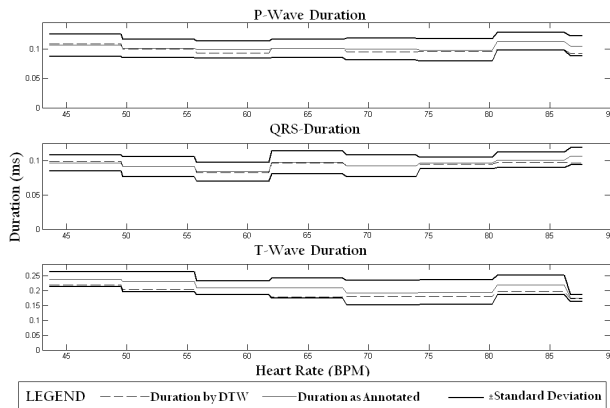


Fig.3: Normal Sinus Constituent Wave Durations as detected by DTW.

constituent waveform for the signals in the database.

Similar results for the Normal Sinus Rhythm test signals are shown in Fig. 3. Observe how the durations identified using the DTW process lie within the annotated variations for each constituent waveform in each heart rate interval.

V. CONCLUSION

One of the most commonly used pattern recognition methods in ECG analysis is DTW. There are a significant number of variations of the DTW algorithm, but many of even the most complex variations of the algorithm use a single difference measurement to classify the best match between an unknown query and a reference signal.

Alternative methods of classification have been suggested that combine DTW with other analysis techniques to further increase the accuracy of the process. The authors have suggested a classifier that uses a time domain based combination of features including heart rate, degree of time alignment required and an amplitude difference measurement. In doing so the DTW method is still used to warp and time align the two signal frames as accurately as possible, but the reference providing the best match is chosen using a number of features from the query and reference signal rather than just a single measurement.

The increased accuracy due to the addition of each feature in the classifier was demonstrated. The contribution of each feature to the classifier was optimized to achieve minimum error by weighting each feature's contribution using scalar coefficients. The performance of the DTW and classifier process was verified by comparison with expert identification of diagnostically significant information from signals within the reference database.

REFERENCES

- [1] H. Sakoe and Seibi Chiba, "Dynamic Programming Algorithm Optimization for Spoken Word Recognition", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-26, No. 1, pp 43-49, February 1978.
- [2] H.J.L.M Vullings, M.H.G. Verhaegen, H.B. Verbruggen, "Automated ECG segmentation with Dynamic Time Warping", *20th Annual Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 20, 1998.
- [3] B. Huang and W. Kinser, ECG Frame Classification Using Dynamic Time Warping, *IEEE Canadian Conference on Electrical & Computer Engineering*, Vol. 2, pp 1105-1110, 2002.
- [4] Venkatesh N and Srinivasan Jayaraman, "Human Electrocardiogram for Biometrics using DTW and FLDA", *The 20th International Conference on Pattern Recognition*, pp 3838-3841, 2010.
- [5] Ali Zifan, Sohrab Saberi, Mohammad Hassan Moradi and Farzad Towhidkhal, "Automated ECG Segmentation Using Piecewise Derivative Dynamic Time Warping", *International Journal of Biological and Life Sciences*, Vol. 1, No. 3, 2005, pp. 181-185.
- [6] Physiobank – Physiological Signal Archives, The ECG Signal Archive, The MIT-BIH QT Database, Available [Online]: <http://www.physionet.org/physiobank/database>.
- [7] R. Jané, A. Blasi, J. García and P. Laguna, Evaluation of an Automatic Threshold Based Detector of Waveform Limits in Holter ECG with the QT Database, *Computer in Cardiology*, Vol. 24, 1997, pp 295-299.
- [8] J.R. Hampton, *The ECG in Practice Third Edition*, Churchill Livingstone, London, 1997, pp. 227.