

# Electrocardiogram Synthesis Using a Gaussian Combination Model (GCM)

S Parvaneh<sup>1,2</sup>, M Pashna<sup>1</sup>

<sup>1</sup>Islamic Azad University, Science and Research Branch, Tehran, Iran

<sup>2</sup>Young Researchers Club, Tehran, Iran

## Abstract

*In this paper modifications to an algorithm for electrocardiogram (ECG) synthesis based on a combination of Gaussians to fit real ECG data have been proposed.*

*A method is proposed for fitting algorithm assuming that constituent Gaussian functions in GCM model are independent. Desired period(s) of ECG were selected and the number of Gaussians in the morphologic model was determined. For ECG synthesis, a Gaussian was fitted around each of the extrema and minimized local error that is defined as local difference of real ECG and our model. The range of Gaussian fitting (place to put independent Gaussian) was determined using two methods: zero crossing method and minimum bank method. Results were presented based on the efficiency of determining the Gaussian parameters in terms of time for fitting and accuracy of model.*

## 1. Introduction

Modeling of ECG can be used in many aspects such as classification (diagnosis), simulators and compression.

Since Gaussian pulse has a shape resembling that of the pulsatile component in ECG, Suppappola et al decomposed and represented ECG as sum of Gaussian pulses but did not account for asymmetries waves in ECG [5]. They proposed Chip Away Decomposition (ChAD) algorithm which was an iterative method for parameter determination. This reference has used normalized root mean square error (NRMSE) as cost function in fitting part of their algorithm. Newton-Raphson, steepest descent and Nelder-Mead Simplex methods have been utilized as their minimization part of the proposed algorithm. They fitted each wave independently and iteratively in their ChAD method.

References [3, 4] improved the proposed model in Reference [1] with accounting for T wave asymmetry. They used 18 parameters (6 Gaussians) in their model and determined them based on a nonlinear least squares optimization procedure in fitting procedure. Modeling of

ECG with 7 Gaussian (21 parameters) have been investigated by Clifford et al [5]. They characterized each of symmetric turning points (Q, R and S) by one Gaussian and asymmetric turning points (P and Q) by two Gaussians. Again a nonlinear gradient descent is performed to optimize parameters of their model.

References [3, 4, and 5] have used beat-by-beat basis in order to derive their model parameters. In their method, all parameters were determined simultaneously. These references also used a simple peak detection and time-aligned averaging to minimize the search space for fitting the parameters. Peak and through detection have also been used to find the relative location of turning points in time.

Our proposed algorithm can be simply used for modeling of one or more cycle(s) without any modification. It also doesn't need any segmentation or windowing to separate ECG cycles and actually all cycles of ECG are modelled together. By zero crossing and minimum bank methods, model calculation time improved without any prior knowledge about place of PQRST waves. In our algorithm, modeling of the main shape of ECG (without accounting noise effect that is important in simulators for ECG training program) results in minimum number of Gaussians. For example if there is no T wave in ECG Waveform, our proposed algorithm will not assign any parameters to T wave. This aspect can be useful in ECG compression.

In this paper effect of

- (a) Two ranges of fitting determination
- (b) Number of Gaussians

on accuracy of ECG modelling and model calculation time is examined.

## 2. Methods

### 2.1. Gaussian function and its role in ECG synthesis

A 1D Gaussian function defined as below:

$$g_{(\mu,\sigma)}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1)$$

In the above equation, mean ( $\mu$ ) and variance ( $\sigma^2$ ) are parameters that identify shape of Gaussian. The Gaussian Combination (GC) is defined as below:

$$GC = \sum_{k=1}^P w_k \cdot g(\mu_k, \sigma_k)(x) \quad (2)$$

Number of Gaussians of GC model in different references is summarized in Table 1.

Table 1. Number of Gaussians in different references

Reference	Number of Gaussians (P)
2	Based on NRMSE threshold
3,4	6
5	7
4	Proposed adaptive determination for $p = n+2m$ ( $n =$ Symmetric turning point, $m =$ asymmetric turning point)
Our Method	Manually or Automatically

## 2.2. Determination of GCM parameters for ECG modeling

In our proposed algorithm, we assumed that ECG is composed of a set of independent Gaussian Functions; so discovery of each independent component place was very important. Steps that should be completed in our algorithm for GCM modeling are listed below:

- 1- Loading ECG signal from Database.
- 2- Selecting cycle(s) of ECG that contain P, QRS and T waves.
- 3- Determining number of Gaussians that are needed in synthesis process. (Further details provided in subsection 2-2-1).
- 4 Applying suitable operation(s) as below, Based on Fitting range determination method :
  - In case of utilizing zero crossing method for fitting range discovery, full rectification of ECG signal (calculating absolute value of ECG signal), separately modeling positive and negative parts of ECG and finally combining Gaussians that have been resulted from that parts should be applied respectively. In order to improve modeling accuracy, positive and negative parts are modelled independently.
  - There is no need to apply any changes to ECG signal for minimum bank method fitting range determination.
- 4- Finding biggest local maximum as Gaussian fitting locus.
- 5- Determining domain of Gaussian fitting around local maximum found in step 5 with zero crossing or minimum bank method (see subsection 2-2-2).
- 6- Fitting a Gaussian function in the part of ECG that has been selected based on a range of fitting found in step 6 (utilize training algorithm that presented in subsection 2-2-3).

- 7- Subtracting the resulting Gaussian function (result of step 7) from original ECG signal in order to determine local error.

$$ECG_{New} = ECG_{Old} - Gaussian_{model} \quad (3)$$

In equation (3)  $ECG_{New}$  is local error that will be used in next step,  $ECG_{Old}$  is remaining parts of ECG from last steps and  $Gaussian_{Model}$  is Gaussian function that have been fitted in step 7 (above).

- 8- Substituting the original signal with local error ( $ECG_{New}$ ) that has been determined in step 8 and back to step 4 until reaching maximum number of iteration (N) or satisfying desired minimum error requirement (iterative process)

### 2.2.1. Determining number of Gaussians in GCM model

There are two methods for determining number of Gaussians. These methods are explained below with their main advantage(s) and disadvantage(s):

- 1- Manual method: Operator will offer number of Gaussians in this. Since calculation duration of synthesis process and memory space needed for GCM model are in direct relation with the number of Gaussians, operator is able to control them (benefit). As this method needs to interact with the operator, it is semi-automatic (drawback).
- 2- Automatic method: Quantity of Gaussians is determined based on number of local maximum that is presented in real ECG. Suitable accuracy in ECG synthesis is the main benefit of this method that corresponds to increase in ECG synthesis time and memory space that is needed. In this method, details of ECG can be included in GC model. This mode is suitable for hardware simulators that can be used for training purpose.

### 2.2.2. Fitting range determination in GCM model

Assigning the range which we decide to fit to an independent Gaussian has great importance in this algorithm. We called this process as fitting range determination and introduced two methods in this work. DC value of ECG should be removed from original signal in both methods. Two methods of range determination in GCM model are:

- 1- Zero crossing method: In this case, two zero crosses were chosen across the biggest local maximum found in step 5 of section 2-2 as starting and ending points of interval for fitting in that range.

- 2- Minimum bank method: In this method, a bank of local minimums was produced. In the next step, a local maximum (biggest local maximum in step 5 of algorithm stated in section 2-2) and two surrounding local minimums formed the fitting interval. These local minimums were found based on the minimum bank gathered previously.

### 2.2.3. Training Algorithm for Gaussian fitting

At this level, a Gaussian function was fitted to the original signal in place of local maximum and the interval around it. In fact, Place of Gaussian (Gaussian mean), Gaussian standard deviation (indicator of Gaussian width) and amplitude of Gaussian in training step should be decided about after determining number of Gaussians. For Gaussian fitting in each interval, the following training method was used:

- 1- Considering place of local maximum as mean of Gaussian function ( $\mu_k$ ).
- 2- Assuming primary guess for  $\sigma_k$ .
- 3- Calculating value of  $w_k$  based on  $\sigma_k$  and signal value in place of Gaussian  $S(x)$ :

$$w_k = \sqrt{2\pi} \cdot S(x) \cdot \sigma_k \quad (4)$$

- 4- Calculating fitting error in desired range as:
 
$$\text{Fitting\_err} = \text{ECG}_{\text{range}} - \text{Gaussian}_{\text{model}} \quad (5)$$

$\text{ECG}_{\text{range}}$  is the part of original ECG located in fitting interval and  $\text{Gaussian}_{\text{model}}$  is as before.

- 5- Computing correction coefficient as below:

$$CF = \frac{\text{Fitting\_err}}{|\text{Fitting\_err}|} \cdot \frac{1}{K} \quad (6)$$

In above Equation CF is correction factor, Fitting\_err is fitting error that determined in step 4 and K is training algorithm step.

- 6- Based on corrected  $\sigma_k$  value returning to step 2 until desired fitting error is satisfied or maximum number of iteration in training step is reached.

## 3. Results

In this section some results of our proposed algorithm in Gaussian Combination Model for ECG modeling with benefits and drawbacks of two proposed fitting range determination methods have been provided. In order to compare the results of our methods, we defined a unique criterion. We described error criterion as area under absolute of local error for this purpose.

A standard ECG with its GCM model (provided by zero crossing method) has been shown in Figure 1.

Using Zero crossing method for fitting interval determination in ECG synthesis for model shown in Figure 1, results in greater error around zero value points rather than other places. Elevation of ECG from zero baseline is also another error rising factor in this method.

Utilizing minimum bank method for discovery of Gaussian fitting interval, as shown in Figure 2, results in lesser errors rather than the above method. In fact, minimum bank method for GCM model and original ECG conform in superior manner. Error in time around 100s arises from non-symmetric shape of T wave. References [3, 4, and 5] proposed usage of two Gaussians for improving these types of errors (this correction needs 6 parameters in order to account for each non-symmetric feature of ECG). Alternatively, we can suggest using of Asymmetrical Gaussian function for compensating errors in asymmetric parts of ECG (5 parameters are needed for each asymmetric feature of ECG). Observations show that minimum bank method performs a better modelling for ECGs which have baseline wandering.

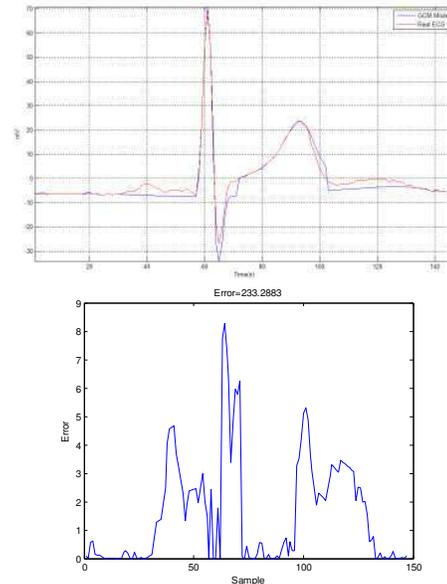


Figure 1. Left: One cycles of standard ECG and its GCM model (interval fitting determination: zero crossing method). Right: Error diagram (error criterion = 233.2).

For comparison of model determination time, we used a computer with Intel Pentium 1.73GHz CPU and 512MB of RAM. In table 2 with variation of Gaussian numbers, we could compare errors of minimum bank and zero crossing methods regarded to each other. This table also provides modeling time needed.

With examination of Table 1, we can conclude that increase in Gaussian numbers improve accuracy of the

model but result in increase of model calculation time. As shown for all different Gaussian numbers, error in minimum bank method is lesser than zero crossing method, but model calculation time is better for the zero crossing method than the minimum bank method.

Also increasing the number of Gaussians (N) has improved accuracy of the synthesis system. With addition to the number of Gaussians, there are disadvantages in increasing model calculation time. It should be noted that increasing number of Gaussians from a particular value will not reduce the error. These errors arise from asymmetrical parts of ECG which Gaussians can't fit.

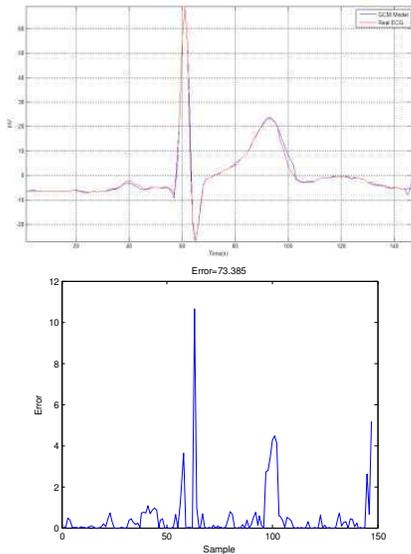


Figure 2. Left: One cycles of standard ECG and its GCM model (interval fitting determination: minimum bank method). Right: Error diagram (error criterion = 73.38).

Table 2. Comparison of errors and algorithm running time of zero crossing and minimum bank methods

minimum bank method		zero crossing method		N
Running Time	Error	Running Time	Error	
0.03s	342.4	0.04s	283.7	4
0.19s	194.1	0.09s	234.8	10
0.44s	102.3	0.23s	233.4	20
1.47s	74.4	0.35s	233.2	50
3.5s	73.3	0.38s	233.2	100
4.8s	73.3	0.42 s	233.2	133

#### 4. Discussion and conclusions

In GCM method, cycle(s) of ECG are represented with combination of Gaussians. In this paper, we assumed that Gaussians in GCM are independent. We used zero

crossing and minimum bank methods in order to find the best place for each independent Gaussian.

Using fitting range determination, fitting method and similarity of morphologic shape of ECG with Gaussian Combination Model provides reasonable results with negligible error.

Zero crossing method is faster than minimum bank method but has lesser accuracy. Modeling time to produce each cycle of ECG is in order of seconds and the number of Gaussians will affect this time.

In this paper attempts were made to provide an algorithm that models the time domain ECG signal directly and based on morphological features. This synthesis method has been used to generate a database of normal and abnormal ECGs that can be used for teaching purposes and evaluation of ECG signal processing algorithms. Each Signal in the database is represented with mean and variance of Gaussians that generate a signal with minimum error. As a result, we have a collection of Gaussian parameters (equal to number of Gaussians that have been used in GCM model).

#### Acknowledgement

Many thanks to Dr. G Clifford for his helpful advice in the preparation of this manuscript. Also we appreciate Eng. Hedyeh K. Mohseni for her support.

#### References

- [1] MsCharry PE, Clifford GD, Tarassenko L, Smith LA. A Dynamical Model For Generating Synthetic Electrocardiogram Signals. *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*. 2003;50: 289–294
- [2] Suppappola S, Sun1 Y, Chiamida SA. Gaussian pulse decomposition: An intuitive model of electrocardiogram waveforms. *Annals of Biomedical Engineering*. 1997;25:252-260
- [3] Clifford GD, Shoeb A, McSharry PE, Janz BA. Model-based filtering, compression and classification of the ECG. *International Journal of Bioelectromagnetism*. 2005;7:158-161
- [4] Clifford GD. A Novel Framework for Signal Representation and Source Separation: Applications to Filtering and Segmentation of Biosignals, *Journal of Biological Systems*. 2006;14:169-183
- [5] Clifford GD, Villarroel M. Model-Based Determination of QT Intervals. *Computers in Cardiology*. 2006; 33:357-360; IEEE Computer Society Press

Address for correspondence

Saman Parvaneh

Iran, Tehran

Sardare Jangal Ave., Golzare1 Ave.

Islamic Azad University, Science and research Branch

School of Biomedical Engineering

E-mail address: saman.parvaneh@gmail.com

Website: www.saman-parvaneh.com