PPGSynth: An innovative toolbox for synthesizing regular and irregular photoplethysmography waveforms

Tang, Q., Chen, Z., Allen, J., Alian, A., Menin, C., Ward, R. & Elgendi, M.

Original citation:
Tang, Q, Chen, Z, Allen, J, Alian, A, Menin, C, Ward, R & Elgendi, M 2020, 'PPGSynth: An innovative toolbox for synthesizing regular and irregular photoplethysmography waveforms', Frontiers in Medicine, vol. 7, 597774.
https://dx.doi.org/10.3389/fmed.2020.597774

DOI 10.3389/fmed.2020.597774
ESSN 2296-858X

Publisher: Frontiers Media

PPGSynth ：用于合成规则和不规则光电体积描记波形的创新工具箱

唐，Q。陈，Z。艾伦，J。Alian，A.，梅恩，C。沃德，R。& Eldell，M。

原文引用：
Tang，Q，Chen，Z，艾伦，J，Alian，A，Menin，C，Ward，R & Elgendi，M 2020，“PPGSynth ：用于合成规则和不规则光电容积描记波形的创新工具箱”，医学前沿，第7卷，597774 。https :// dx.doi .org / 10.3389 / fmed .2020.597774

DOI 10.3389 / fmed .2020.597774
ESN 2296 - 858X

出版社：Frontiers Media

frontiers in Medicine

TECHNOLOGY AND CODE
published: 02 November 2020
doi: 10.3389/fmed.2020.597774

PPGSynth: An Innovative Toolbox for Synthesizing Regular and Irregular Photoplethysmography Waveforms

Qunfeng Tang [1,2], Zhencheng Chen [2], John Allen [3,4], Aymen Alian [5], Carlo Menon [6], Rabab Ward [1] and Mohamed Elgendi [1,6,7,8]*

[1] Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada, [2] School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin, China, [3] Faculty of Medical Sciences, Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, United Kingdom, [4] Research Centre for Intelligent Healthcare, Coventry University, Coventry, United Kingdom, [5] Yale School of Medicine, Yale University, New Haven, CT, United States, [6] School of Mechatronic Systems Engineering, Simon Fraser University, Burnaby, BC, Canada, [7] Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada, [8] BC Children's & Women's Hospital, Vancouver, BC, Canada

Photoplethysmography (PPG) is increasingly used in digital health, exceptionally in smartwatches. The PPG signal contains valuable information about heart activity, and there is lots of research interest in its means and analysis for cardiovascular diseases. Unfortunately, to our knowledge, there is no arrhythmic PPG dataset publicly available—this paper attempt to provide a toolbox that can generate synthesized arrhythmic PPG signals. The model of a single PPG pulse in this toolbox utilizes two combined Gaussian functions. This toolbox supports synthesizing PPG waveform with regular heartbeats and three irregular heartbeats: compensation, interpolation, and reset. The user can generate a large amount of PPG data with a certain irregularity, with different sampling frequency, time length, and a range of noise types (Gaussian noise and multi-frequency noise) can be added to the synthesized PPG which can all be modified from the interface, and different types of arrhythmic PPGs (as calculated by the model) generated. The generation for large PPG datasets that simulate PPG collected from real humans could be used for testing the robustness of developed algorithms that are targeting arrhythmic PPG signals. Our PPG synthesis tool is publicly available.

Keywords: digital health, data modeling, data generation, big data, biosignal generation, PPG construction, signal simulation, generative model

OPEN ACCESS

Edited by:
Ovidiu Constantin Baltatu, Khalifa University, United Arab Emirates

Reviewed by:
Paolo Crippa, Marche Polytechnic University, Italy
Aniruddha Sinha, Tata Consultancy Services, India

****Correspondence:***
Mohamed Elgendi
moe.elgendi@gmail.com

Specialty section:
This article was submitted to Translational Medicine, a section of the journal Frontiers in Medicine

Received: *22 August 2020*
Accepted: *01 October 2020*
Published: *02 November 2020*

Citation:
Tang Q, Chen Z, Allen J, Alian A, Menon C, Ward R and Elgendi M (2020) PPGSynth: An Innovative Toolbox for Synthesizing Regular and Irregular Photoplethysmography Waveforms. Front. Med. 7:597774. doi: 10.3389/fmed.2020.597774

1. INTRODUCTION

Photoplethysmogram (PPG) signal contains rich information about the cardiovascular system (1). In the past decade, some studies have used PPG to calculate heart rate, oxygen saturation, blood pressure, cardiac output, cardiac index, peripheral vascular resistance, and other indicators of cardiovascular function, and many algorithms were developed to calculate these indices (2).

Four PPG databases, at time of writing, are publicly available: Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) (3), the University of Queensland Vital Signs Dataset (4), Vortal Dataset (5), and PPG-BP (6). The sampling frequency and time length of PPG signals are different in different databases; however, most algorithms designed for these databases are signal-independent. Additionally, it is still a challenge to evaluate the performance of these algorithms under different PPG types and different signal-to-noise ratios (SNR).

技术和代码
发布：02十一月2020
doi：10.3389/fmed.2020.597774

PPGSynth：一种用于合成规则和不规则光电容积描记波形的创新工具

Qunfeng Tang、Zhencheng Chen、John Allen、Aymen Alian、Carlo Menon、Rabab Ward 和Mohamed Elgendi*

1加拿大温哥华不列颠哥伦比亚省大学电气与计算机工程系、中国桂林桂林电子科技大学电子工程与自动化学院、英国泰恩河畔纽卡斯尔纽卡斯尔大学医学院人口健康科学研究所、英国考文垂考文垂大学智能医疗研究中心美国康涅狄格州纽黑文市耶鲁大学耶鲁医学院、加拿大卑诗省伯纳比市西蒙弗雷泽大学机电系统工程学院、加拿大卑诗省温哥华市不列颠哥伦比亚省大学医学院、加拿大卑诗省温哥华市卑诗儿童妇女医院

光电体积描记术（PPG）越来越多地用于数字健康，特别是智能手表。PPG信号包含有关于心脏活动的有价值的信息，并且在心血管疾病的手段和分析方面有很多研究兴趣。不幸的是，据我们所知，还没有公开的心电PPG数据集-本文试图提供一个工具箱，可以生成合成的心电PPG信号。该工具箱中的单个PPG脉冲的模型利用两个组合的高斯函数。该工具箱支持合成PPG波形，包括规则心跳和三种不规则心跳：补偿、插值和重置。用户可以生成具有一定不规则性的大量PPG数据，具有不同的采样频率、时间长度，并且可以将一系列噪声类型（高斯噪声和多频噪声）添加到合成PPG，这些都可以从界面进行修改，并且生成不同类型的动态PPG（如由模型计算的）。模拟从真实的人收集的PPG的大型PPG数据集的生成可以用于测试以运动PPG信号为目标的所开发的算法的鲁棒性。我们的PPG合成工具是公开的。

OPEN ACCESS

编辑：
Ovidiu Constantin Baltatu，
哈利法大学，
阿拉伯联合酋长国

复核人：
意大利马尔凯理工大学Paolo Crippa
Aniruddha Sinha，印度塔塔咨询服务公司

*通信：
穆罕默德·埃尔根迪
moeelgendi@gmail.com

专业部分：
本文提交给
转化医学，
杂志的一部分
医学前沿

投稿时间：2020年8月22日
接受日期：2020年10月1日
发布日期：2020年11月2日

引文：
Tang Q、Chen Z、Allen J、Alian A、Menon C、Ward R和Elgendi M（2020年）PPGSynth：合成常规和合成的创新工具箱

不规则光电体积描记波形。前面7：597774

doi：10.3389/fmed.2020.597774

关键词：数字健康，数据建模，数据生成，大数据，生物信号生成，PPG构建，信号模拟，生成模型

1.介绍

光电容积描记图（PPG）信号包含关于心血管系统的丰富信息（1）。在过去的十年中，一些研究已经使用PPG来计算心率、血氧饱和度、血压、心输出量、心脏指数、外周血管阻力和其他心血管功能指标，并开发了许多算法来计算这些指标（2）。

在撰写本文时，有四个PPG数据库是公开的：重症监护多参数智能监测（MIMIC）（3）、昆士兰州大学生命体征数据集（4）、Vortal数据集（5）和PPG-BP（6）。PPG信号的采样频率和时间长度在不同的数据库中是不同的;然而，为这些数据库设计的大多数算法是与信号无关的。此外，在不同PPG类型和不同信噪比（SNR）下评估这些算法的性能仍然是一个挑战。

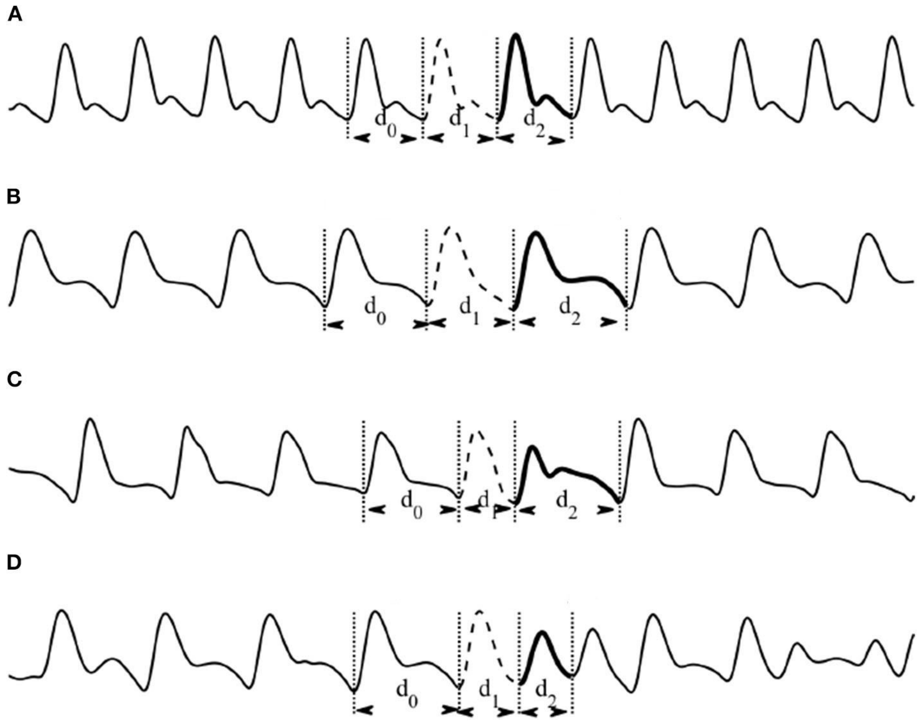

FIGURE 1 | Types of heartbeat classification based sinus node response to atrial premature depolarization [adapted from (7, 8)]. **(A)** Regular: $d_0 = d_1 = d_2$. **(B)** Compensation: $d_1 + d_2 = 2d_0$. **(C)** Reset: $d_0 < d_1 + d_2 < 2d_0$. **(D)** Interpolation: $d_1 + d_2 = d_0$.

PPGSynth is developed to generate PPG signals across a wider range of sampling frequencies and time lengths. Three types of irregular PPG signals also can be generated by the PPGSynth tool. It can also conveniently manage parameters and graphical output through a graphical user interface (GUI). This toolbox does not require highly experienced users, but it is recommended that you have basic knowledge of PPG signal and cardiac irregularities.

2. HEARTBEAT CLASSIFICATION

The amplitude, duration, and waveform shape of PPG pulses tend to vary between persons, and they even differ from moment to moment in the same person. Premature heartbeats are typical irregular PPG beats. There are two different types of premature heartbeats, premature atrial contractions and premature ventricular contractions. This study only focuses on irregular PPG signals that have premature atrial contractions. Premature atrial contraction changes the waveform of PPG for two consecutive beats. In this study, these two beats are defined as the premature group, and the first beat of the premature group and the second beat of the premature group are defined as the first beat and second beat, respectively. The beats without the influence of premature contractions are defined as reference beats. The first beat duration is always less than the reference beat duration. Based on the difference between the durations of the first beat and second beat, Roskamm and Csapo (7), classified heartbeats into four types: compensation, reset, interpolation, and re-entry. Based on their analysis, these four types are defined as follows:

- Compensation: the second beat is prolonged, and the sum of the first beat duration and second beat duration is equal to the duration of two reference beats.
- Reset: the second beat is prolonged, but the sum of the first beat duration and second beat duration is less than the duration of two reference beats.
- Interpolation: the sum of the first beat duration and second beat duration is equal to one reference beat duration.
- Re-entry: the sum of the first beat duration and second beat duration is less than one reference beat duration. We could not find a template, within the four databases mentioned above, that satisfies the definition of re-entry. Therefore, the re-entry is not included in the current analysis.

A previous attempt (8) on the use of heartbeat classification using ECG signals inspired the classification of heartbeats in PPG signals. Based on the previous heartbeat classification (8), **Figure 1A** shows the regular heartbeats where the first beat, second beat, and third beat have equal durations (e.g., $d_0 = d_1 = d_2 = 1,000$ ms). On the other hand, **Figure 1B** shows the compensation phase, the second beat (e.g., $d_1 = 850$ ms) is followed by a prolonged beat (e.g., $d_2 = 1150$ ms) to compensate the two beats duration of 2,000 ms. During the reset phase (**Figure 1C**), the second beat (e.g., $d_1 = 650$ ms) is followed by a prolonged beat (e.g., 1,150 ms), while in the interpolation

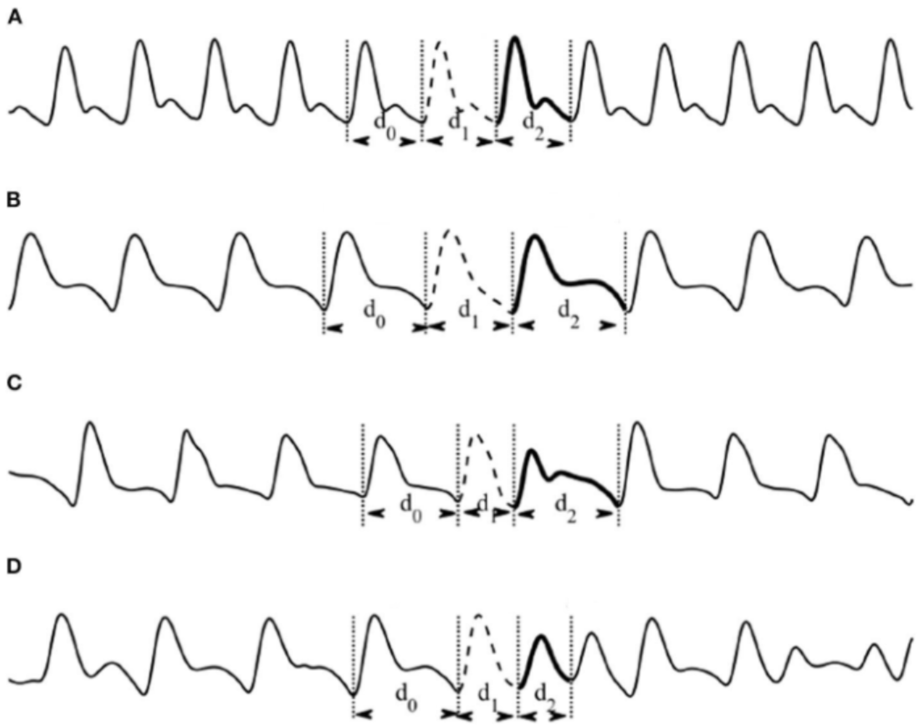

图1| 基于心跳分类的窦房结对房性早搏除极反应的类型[改编自（7，8）]。(A)常规：d=d=d。(B)补偿：d+d=2d。(C)重置：d<d+d<2d。(D)插值：d+d=d。

开发PPGSynth 以在更宽的采样频率和时间长度范围内生成PPG 信号。PPGSynth 工具还可以生成三种类型的不规则PPG 信号。它还可以通过图形用户界面（GUI）方便地管理参数和图形输出。此工具箱不需要经验丰富的用户，但建议您具备PPG 信号和心脏不规则性的基本知识。

2.心跳分类

PPG 脉冲的幅度、持续时间和波形形状往往在人与人之间变化，并且它们甚至在同一个人中时刻与时刻不同。早搏是典型的不规则PPG 搏动。有两种不同类型的早搏，房性早搏和室性早搏。本研究仅关注具有房性早搏的不规则PPG 信号。房性期前收缩改变连续两个心搏的PPG 波形。在本研究中，将这两个心搏定义为早搏组，将早搏组的第一个心搏和早搏组的第二个心搏分别定义为第一心搏和第二心搏。不受过早收缩影响的心搏被定义为参考心搏。第一心跳持续时间总是小于参考心跳持续时间。根据第一次心跳和第二次心跳的持续时间之间的差异，Roskamm 和Csapo （7）将心跳分为四种类型：补偿，重置，插值，和重返大气层根据他们的分析，这四种类型的定义如下：

- 补偿：第二心跳被延长，并且第一心跳持续时间和第二心跳持续时间之和等于两个参考心跳的持续时间。
- 重置：第二个心跳被延长，但第一个心跳持续时间和第二个心跳持续时间之和小于两个参考心跳的持续时间。
- 插值：第一拍持续时间和第二拍持续时间之和等于一个参考拍持续时间。
- 重新输入：第一次搏动持续时间和第二次搏动持续时间之和小于一个参考搏动持续时间。在上述四个数据库中，我们找不到符合重返定义的模板。因此，本分析不包括重入。

关于使用ECG 信号的心跳分类的使用的先前尝试（8）启发了PPG 信号中的心跳的分类。基于先前的心跳分类（8），图1A示出了规则心跳，其中第一心跳、第二心跳和第三心跳具有相等的持续时间（例如，d=d=d=1，000 ms）。另一方面，图1B示出了补偿阶段，第二拍（例如，D=850 ms）之后是延长的搏动（例如，d= 1150ms）以补偿2，000 ms 的两个搏动持续时间。在复位阶段（图1C）期间，第二搏动（例如，D=650 ms）之后是延长的搏动（例如，1，150ms），而在插值中

(**Figure 1D**), the second beat (e.g., $d_1 = 400$ ms) is followed by an irregular beat (e.g., $d_2 = 600$ ms).

3. METHODOLOGY

The PPGSynth consists of three main parts: the model of a single PPG pulse, the pulse duration generator, and the noise generator.

3.1. Model of Single PPG Pulse

The single PPG pulse step is based on a recently published model (9) that simulates fingertip PPG waveforms. Note that the adopted model (9) is an early work on healthy subjects; however, this paper is about arrhythmic PPG beats relating to cardiovascular patient simulated recordings, which is definitely a new concept.

The construction of a PPG waveform is regarded as a motion trajectory in the three-dimensional space established by the coordinate system (x, y, z). As shown in **Figure 2**, the periodicity of PPG is represented by a circular motion.

The motion trajectory in the (x, y) plane is the unit circle. One cycle of movement on the circle corresponds to a peak-to-peak interval or heartbeat. The trajectory in the z direction is the PPG signal. The systolic wave and diastolic wave are simulated in Gaussian functions. The equation of (x, y, z) is defined as follows:

$$\begin{cases} x(t) = \cos(\omega(t - t_0) - \pi) \\ y(t) = \sin(\omega(t - t_0) - \pi) \\ z(t) = \sum_{i=1}^{2} a_i \exp(-\frac{(\theta(t)-\theta_i)^2}{2b_i^2}), \end{cases} \quad (1)$$

where t is time, ω is the angular velocity (which is used to control the pulse duration), t_0 is the end time of the previous beat,π is used to align the initial point of this model to the position of the onset in a PPG waveform, and a_i, θ_i, and b_i are the amplitude of the peak, the position of the center of the peak, and the standard deviation of Gaussian functions, respectively. Additionally, ω is calculated by:

$$\omega = \frac{2\pi}{T}, \quad (2)$$

where T is the PPG pulse duration. θ is the four-quadrant inverse tangent of (x, y), which is introduced as an independent variable for motion in the z direction and is defined as:

$$\theta(t) = atan2(y(t), x(t)), \quad (3)$$

with the changes to (x, y), θ is in the range of $(-\pi, \pi)$.

The corresponding changes to x, y, and z over a single period are shown in **Figure 2**; these are repeated in the next pulses. In this figure, the pulse duration was 1 s, and the sampling frequency was 125 Hz. Obtaining a waveform of the synthetic PPG pulse that is close as possible to the real PPG pulse through calculation of model parameters is an optimization problem (finding the optimal parameters). The objective function was expressed as follows:

$$p^* = \arg\min_p(\sum_{n=1}^{l} (z_p(n) - s(n))^2 + (1 - corr(z_p(n), s(n)))), \quad (4)$$

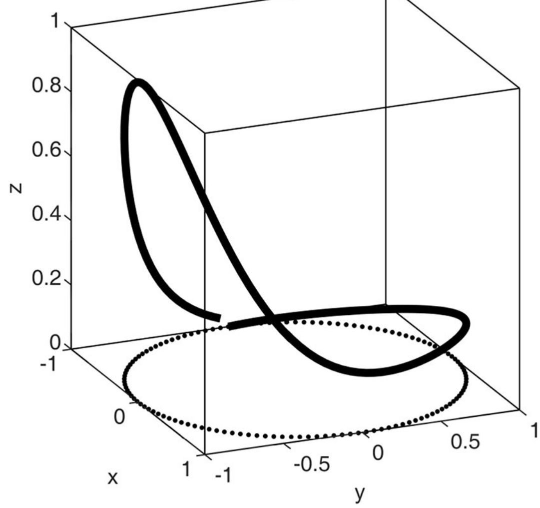

FIGURE 2 | Motion trajectory of a single synthesized PPG waveform. This figure shows how the Gaussian model simulates a heartbeat of a PPG waveform.

such that $p = \{a_1, \theta_1, b_1, a_2, \theta_2, b_2\}$,
with the constraints:

$$\begin{cases} 0 \le a_2 < a_1 \le 1 \\ 0 \le b_1 < b_2 \le 3 \\ -\pi \le \theta_2 < \theta_1 \le \pi, \end{cases} \quad (5)$$

where $z_p(n)$ is the synthetic PPG, l is the length of the real PPG $s(n)$, and $corr$ is Pearson's linear correlation coefficient.

In this study, the interior-point (10) method was used to solve the optimization problem.

3.2. Variability of Parameters

In real-world PPG, the waveform often varies between pulses—sometimes dramatically so. To make the synthesized PPG closer to a real PPG, we used a Gaussian distribution to generate random parameters for our model. In this paper, the mean value and standard deviation of each parameter's Gaussian distribution are derived from real PPG signals, set by modeling a PPG pulse (from start of a pulse to the start of the consecutive pulse). For a PPG trace which has regular beats; we modeled all pulses in a 5-min PPG from the MIMIC database (3). However, for a PPG trace which has irregular beats, the waveform of the first beat and second beat in the premature group is distinct from the reference beat. We model three compensation segments (include the first beat and second beat) from one record of the Queensland database (4) to get the distribution of parameters in the compensation type. For reset, four reset segments from one record of the MIMIC database are used to get the distribution. For interpolation, three interpolation segments in one record from the Queensland database are used to calculate the parameters. The mean and standard deviation of these parameters are shown in **Table 1**.

（图1D），第二次搏动（例如，d=400 ms）之后是不规则的搏动（例如，d=600 ms）。

3.方法

PPGSynth 由三个主要部分组成：单个PPG脉冲的模型、脉冲持续时间发生器和噪声发生器。

3.1单个PPG脉冲的模型

单个PPG脉冲步长基于最近发布的模拟指尖PPG波形的模型（9）。请注意，所采用的模型（9）是对健康受试者的早期研究;然而，本文是关于与心血管患者模拟记录相关的心电PPG搏动，这绝对是一个新概念。

PPG波形的构造被视为由坐标系（x，y，z）建立的三维空间中的运动轨迹。如图2所示，PPG的周期性由圆周运动表示。

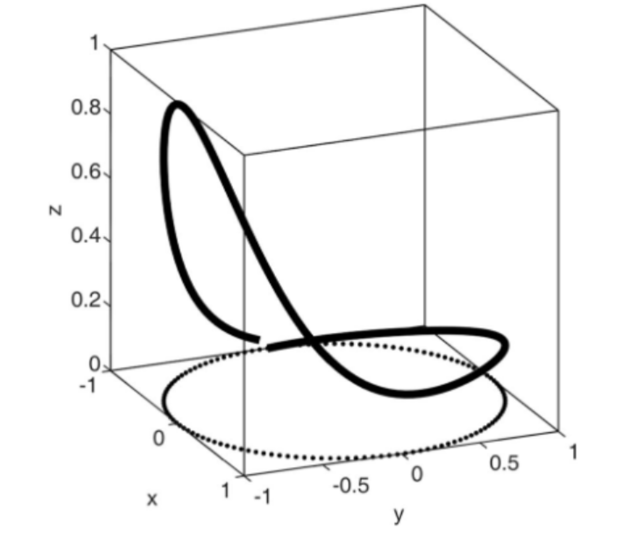

图2| 单个合成PPG波形的运动轨迹。该图示出了高斯模型如何模拟PPG波形的心跳。

（x，y）平面中的运动轨迹是单位圆。圆上的一个运动周期对应于峰-峰间期或心跳。z方向上的轨迹是PPG信号。收缩波和舒张波用高斯函数模拟。（x，y，z）的等式定义如下：

$$x(t) = \cos(\omega(t-t) - \pi)$$
$$y(t) = \sin(\omega(t-t) - \pi) \quad (1)$$
$$z(t) = \sum_{i=1} a \exp\left(-\frac{}{2b}\right),$$

其中t是时间（由d控制脉冲持续时间），是前一次搏动的结束时间，π用于将该模型的初始点与PPG波形中的起始位置对齐，α、θ分别表示峰值的幅度、峰值中心的位置和高斯函数的标准差。此外，ω通过以下公式计算：

$$\omega = \frac{2\pi}{T}, \quad (2)$$

其中T是PPG脉冲持续时间。θ被引入作为z方向上的运动的独立变量，并且被定义为：

$$\theta(t) = \operatorname{atan2}(y(t), x(t)), \quad (3)$$

随着对（x，y）的改变，θ 在（−π，π）的范围内。

在一个周期内x、y和z的相应变化如图2所示;这些变化在下一个脉冲中重复。在该图中，脉冲持续时间为1秒，采样频率为125 Hz。通过模型参数的计算获得尽可能接近真实的PPG脉冲的合成PPG脉冲的波形是优化问题（找到最佳参数）。目标函数表示如下：

$$p = 最小参数_{p} \left(\sum_{n=1}^{l} (z(n) - s(n)) + (1 - corr(z(n), s(n)), \right. \quad (4)$$

使得p={a，θ，B，a，θ，B}，约束条件为：

$$0 \le a < a \le 1$$
$$0 \le b < b \le 3 \quad (5)$$
$$-\pi \le \theta < \theta \le \pi,$$

其中z（n）是合成PPG，l是真实的PPG s（n）的长度，并且corr是皮尔逊线性相关系数。

在本研究中，采用邻点（10）法来解决优化问题。

3.2参数的可变性

在真实世界的PPG中，波形通常在脉冲之间变化- 有时变化很大。为了使合成的PPG更接近于真实的PPG，我们使用高斯分布来为我们的模型生成随机参数。在本文中，每个参数的高斯分布的平均值和标准偏差是从通过对PPG脉冲（从脉冲的开始到连续脉冲的开始）建模而设置的真实的PPG信号导出的。对于具有规则搏动的PPG迹线;我们从MIMIC数据库中建模了5分钟PPG中的所有脉冲（3）。然而，对于具有不规则搏动的PPG迹线，过早组中的第一搏动和第二搏动的波形与参考搏动不同。我们从昆士兰州数据库（4）的一个记录中对三个补偿段（包括第一拍和第二拍）进行建模，以获得补偿类型中的参数分布。对于复位，MIMIC数据库的一个记录中的四个复位段用于获得分布。对于插值，使用来自昆士兰州数据库的一个记录中的三个插值段来计算参数。这些参数的平均值和标准差见表1。

TABLE 1 | The optimal parameters for each PPG template.

PPG type	Beat	a_1	a_2	b_1	b_2	θ_1	θ_2	Ratio of pulse duration
Regular	-	0.997 ± 0.028	0.225 ± 0.030	0.641 ± 0.034	0.937 ± 0.161	-1.471 ± 0.147	1.019 ± 0.102	-
Compensation	*1stBeat*	0.829 ± 0.010	0.420 ± 0.018	0.732 ± 0.033	1.219 ± 0.021	-1.008 ± 0.147	0.450 ± 0.167	0.830 ± 0.061
Compensation	*2ndBeat*	0.785 ± 0.034	0.405 ± 0.049	0.678 ± 0.036	1.115 ± 0.065	-1.792 ± 0.080	-0.607 ± 0.107	1.170 ± 0.061
Reset	*1stBeat*	0.774 ± 0.012	0.774 ± 0.012	0.647 ± 0.041	1.007 ± 0.046	-1.378 ± 0.180	0.173 ± 0.180	0.607 ± 0.019
Reset	*2ndBeat*	0.995 ± 0.002	0.197 ± 0.024	0.778 ± 0.055	1.045 ± 0.341	-1.809 ± 0.203	0.892 ± 0.325	0.596 ± 0.484
Interpolation	*1stBeat*	0.668 ± 0.151	0.490 ± 0.006	0.893 ± 0.034	1.428 ± 0.062	-0.627 ± 0.292	0.442 ± 0.635	0.561 ± 0.028
Interpolation	*2ndBeat*	0.595 ± 0.084	0.537 ± 0.092	0.889 ± 0.170	1.321 ± 0.289	-1.049 ± 0.207	-0.289 ± 0.480	0.475 ± 0.028

Here, θ_1 and θ_2 are the locations of the first Gaussian distribution and the second Gaussian distribution, respectively. Here, a_1 and a_2 are the amplitudes, and, b_1 and b_2 are the widths of the first and second Gaussian distributions, respectively. Note that in the regular PPG heartbeat, all beats have similar durations based on the standard deviation.

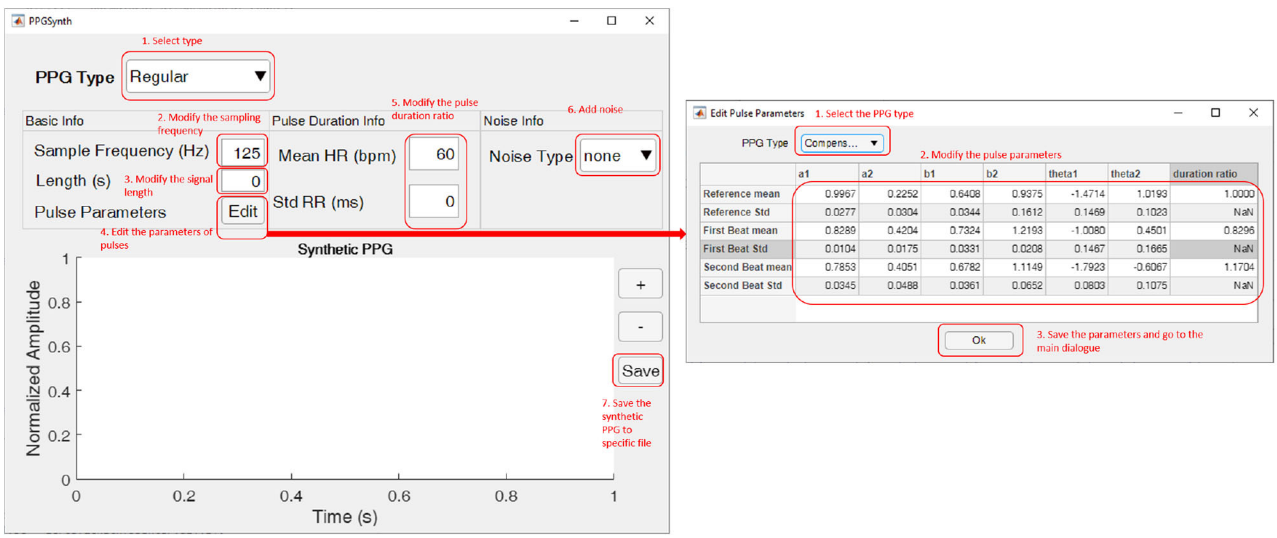

FIGURE 3 | The graphical interface of PPGSynth. The left side is the main dialogue. The edited parameters dialogue on the right side will pop up after pressing the "Edit" button. The red text show the numbered steps for generating PPGs. This figure shows how to use the graphical interface of PPGSynth.

Note that since we do not have a high-quality re-entry PPG in our database, this toolbox does not support generating re-entry PPGs. To not change the irregular category, the duration ratio of irregular beat and regular beat in synthetic PPG uses a fixed value instead of a random number obeying Gaussian distribution. These fixed values are the mean of the ratio of the pulse duration of the irregular beat and regular beat in **Table 1**.

3.3. Pulse Duration

In this study, the PPG pulse duration is defined as the valley-to-valley interval. To generate a sequence of PPGs, a series of PPG pulse durations were needed. In this toolbox, reference pulse durations are generated based on the basic heart rate and signal time lengths, and then the reference pulse durations are randomly replaced by two consecutive irregular beats. The ratios of the first beat duration and second beat duration to the reference beat duration are calculated from each type of PPG templates, and the results are shown in **Table 1**.

3.4. Adding Noise

Two types of noise are available in this toolbox: white Gaussian noise and multi-frequency noise. Multi-frequency noise is a set of noises that have different amplitudes and frequencies. Each noise is generated as follows:

$$n(t) = A \sin 2\pi f t, \tag{6}$$

where A is the amplitude of the peak of the noise and f is the frequency of the noise.

If necessary, users can add one or more different amplitudes and frequency noises to the clean synthetic PPG signals.

4. CUBIC INTERPOLATION

The variability of parameters will make the endpoint value of one beat differ from the next beat's onset value. In this paper, cubic interpolation was used to smooth the synthetic PPG. Cubic spline interpolation involves a spline where each piece is a third-degree polynomial specified by its values and first derivatives at the corresponding domain interval's endpoints. The interpolation

表1| 每个PPG模板的最佳参数。

PPG类型	Beat	a	a	b	b	θ	θ	脉宽比
定期	-	0.997 ± 0.028	0.225 ± 0.030	0.641 ± 0.034	0.937 ± 0.161	-1.471 ± 0.147	1.019 ± 0.102	-
补偿	1stBeat	0.829 ± 0.010	0.420 ± 0.018	0.732 ± 0.033	1.219 ± 0.021	-1.008 ± 0.147	0.450 ± 0.167	0.830 ± 0.061
补偿	2ndBeat	0.785 ± 0.034	0.405 ± 0.049	0.678 ± 0.036	1.115 ± 0.065	-1.792 ± 0.080	-0.607 ± 0.107	1.170 ± 0.061
复位	第一次心跳	0.774 ± 0.012	0.774 ± 0.012	0.647 ± 0.041	1.007 ± 0.046	-1.378 ± 0.180	0.173 ± 0.180	0.607 ± 0.019
复位	第二拍	0.995 ± 0.002	0.197 ± 0.024	0.778 ± 0.055	1.045 ± 0.341	-1.809 ± 0.208	0.892 ± 0.325	0.596 ± 0.484
插值	第一次心跳	0.668 ± 0.151	0.490 ± 0.006	0.893 ± 0.034	1.428 ± 0.062	-0.627 ± 0.292	0.442 ± 0.635	0.561 ± 0.028
插值	第二拍	0.595 ± 0.084	0.537 ± 0.092	0.889 ± 0.170	1.321 ± 0.289	-1.049 ± 0.207	-0.289 ± 0.480	0.475 ± 0.028

这里，θ和θ分别是第一高斯分布和第二高斯分布的位置。这里，*a*和*a*分别是第一高斯分布和第二高斯分布的振幅和带宽。请注意，在常规*PPG*心跳中，所有心跳都具有基于标准差的相似持续时间。

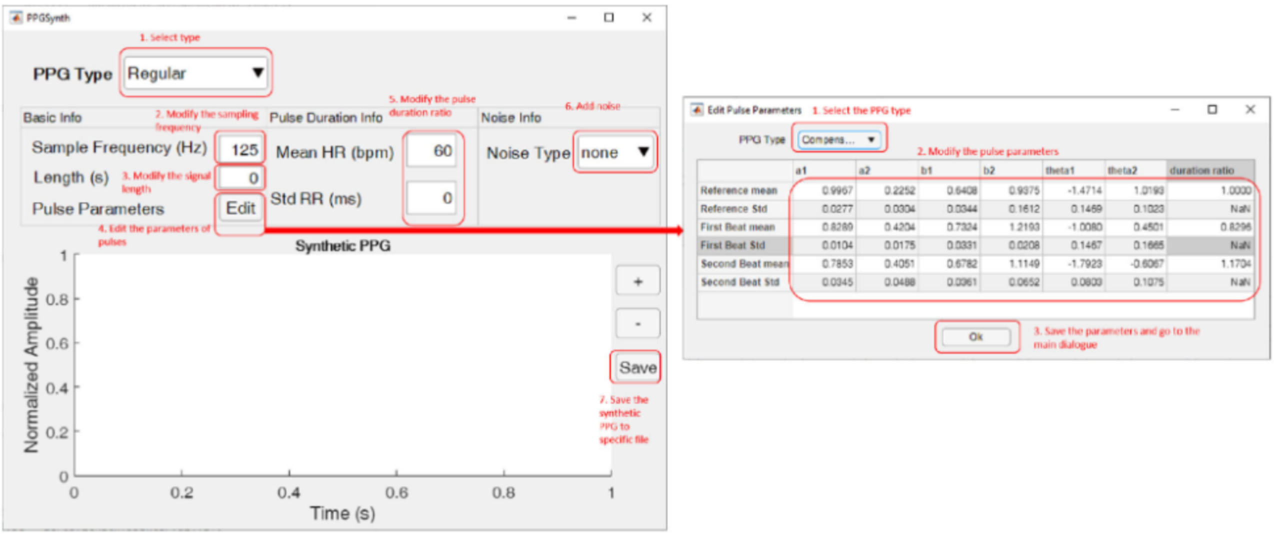

图3| PPGSynth的图形界面。左边是主要的对话。按下"编辑"按钮后，右侧弹出编辑参数对话框。红色文本显示生成PPG的编号步骤。下图显示了如何使用PPGSynth的图形界面。

请注意，由于我们的数据库中没有高质量的再入PPG，因此此工具箱不支持生成再入PPG。为了不改变不规则类别，合成PPG中的不规则搏动和规则搏动的持续时间比使用固定值而不是服从高斯分布的随机数。这些固定值是表1中不规则搏动和规则搏动的脉宽比的平均值。

3.3脉冲持续时间

在本研究中，PPG脉冲持续时间定义为谷-谷间期。为了生成PPG序列，需要一系列PPG脉冲持续时间。在该工具箱中，基于基本心率和信号时间长度生成参考脉冲持续时间，然后将参考脉冲持续时间随机替换为两个连续的不规则心跳。根据每种类型的PPG模板计算第一搏动持续时间和第二搏动持续时间与参考搏动持续时间的比率，并且结果在表1中示出。

3.4添加噪声

此工具箱中提供两种类型的噪声：白色高斯噪声和多频噪声。多频噪声是一组具有不同幅度和频率的噪声。每个噪声的生成如下：

$$n(t) = A \cdot \sin(2\pi f t), \quad (6)$$

其中A是噪声峰值的幅度，f是噪声的频率。

如果需要，用户可以将一个或多个不同的振幅和频率噪声添加到干净的合成PPG信号。

4.三次插值

参数的可变性将使一个心搏的终点值与下一个心搏的起始值不同。本文采用三次插值方法对合成的PPG图像进行平滑处理。三次样条插值涉及样条，其中每一段都是由其值和相应域区间端点处的一阶导数指定的三次多项式。插值

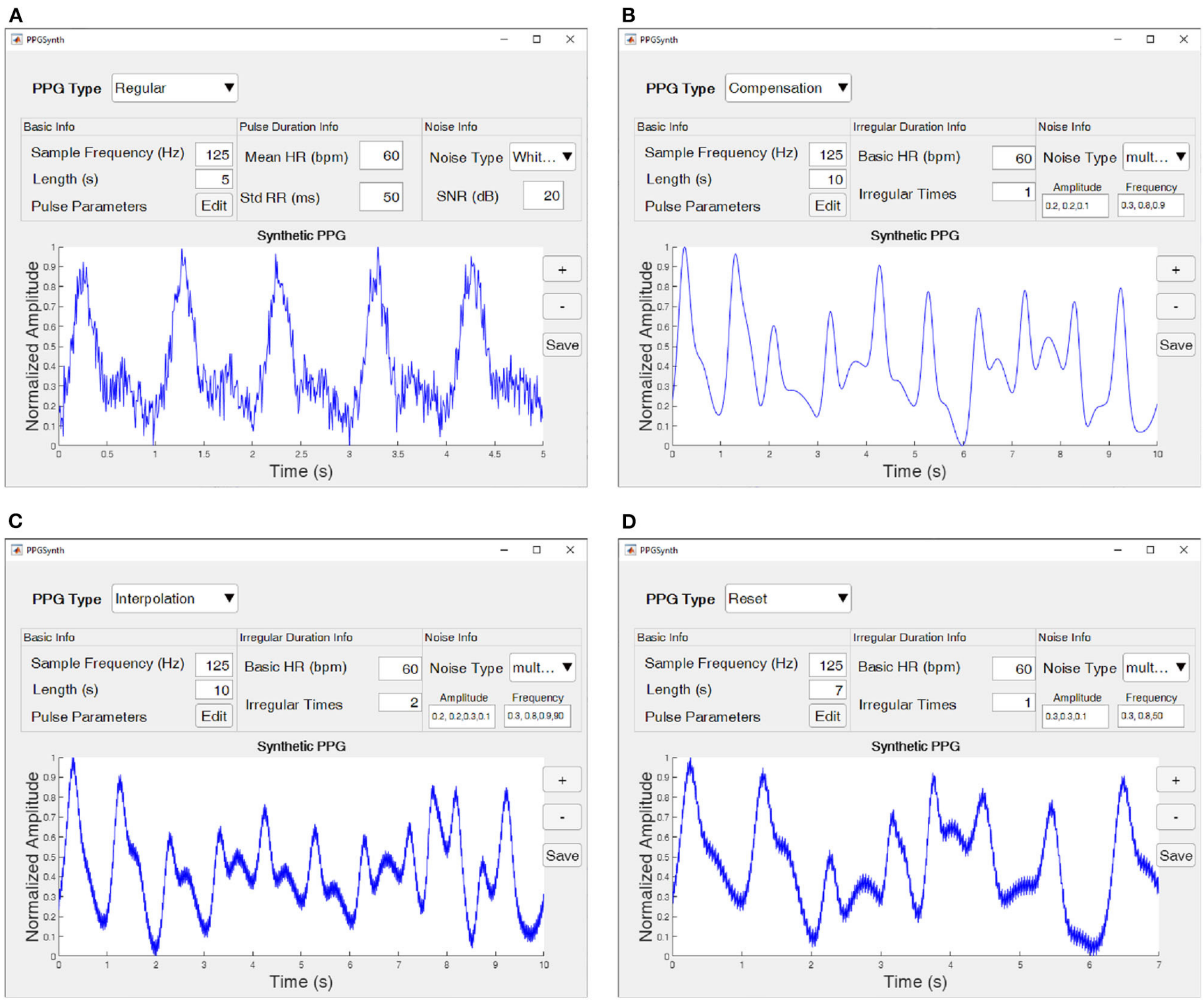

FIGURE 4 | Four types of PPG heartbeat generated using the PPGSynth. **(A)** A 5-s regular PPG with white Gaussian noise. **(B)** A 10 s compensation PPG with multi-frequency noise. **(C)** A 10-s interpolation PPG with multi-frequency noise. **(D)** A 10-s reset PPG with multi-frequency noise. The parameters of each type shown at each subfigure. Additionally, the pulse parameters associated with each type are shown in **Table 1**.

involved a total of 0.2 s around the onset. The previous sampling points of 0.05 s and the last sampling points of 0.05 s were used to fit the interpolation function. The middle 0.1-s samples' value is replaced by the corresponding samples' value generated by the interpolation function.

5. THE GRAPHICAL USER INTERFACE

Figure 3 shows the main dialogue of the GUI. The first step is to select the type of synthetic PPG using the drop-down button in the upper left corner. Available options are regular or three types of irregular PPG. Then we can modify the sampling frequency and signal length in the "Basic Info" panel. Once we change any data, the GUI will attempt to generate the synthetic PPG and show it at the bottom of the dialogue.

By pressing the "Edit" button, users can modify the parameters of pulses in the pop-up dialogue. For a regular PPG, this toolbox uses the same parameters for different pulses. But for irregular PPGs, parameters are different in the first beat of the premature group, the second beat of the premature group, and the reference beat. Users can also modify the ratio of first beat duration and second beat duration to reference beat duration. The default value is shown in **Table 1**. Once done with editing pulse parameters, press the "OK" button to save these parameters and go to the main dialogue.

After setting the basic info, users should set some parameters to generate the pulse duration. For a regular PPG, users can modify the mean heart rate and standard deviation of the RR intervals in the "Pulse Duration Info" panel. For irregular PPG types, this panel changes to an "Irregular Duration Info" panel, where users can modify the basic heart rate and irregular times of

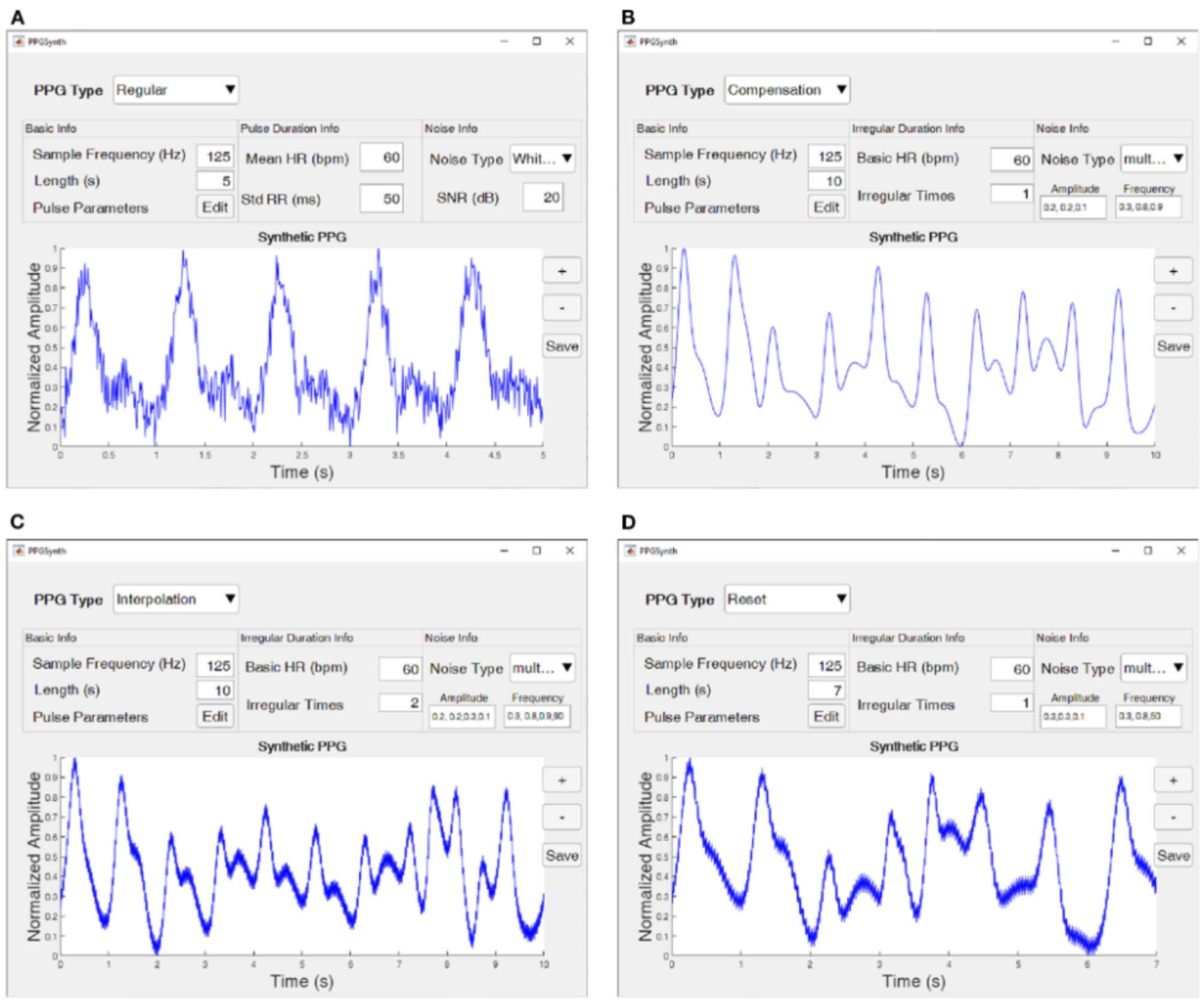

图4| 使用PPGSynth生成的四种类型的PPG心跳。(A)具有白色高斯噪声的5-s规则PPG。(B)具有多频噪声的10-s补偿PPG。(C)具有多频噪声的10-s插值PPG。(D)10-s重置PPG，具有多频噪声。每种类型的参数显示在每个子图中。另外，与每种类型相关联的脉冲参数在表1中示出。

在发病前后总共有0.2 秒。使用0.05 s的先前采样点和0.05 s的最后采样点来拟合插值函数。中间的0.1 - s采样值被插值函数生成的相应采样值替换。

5.图形用户界面

图3 显示了GUI的主对话框。第一步是使用左上角的下拉按钮选择合成PPG 的类型。可用选项是常规或三种类型的不规则PPG 。然后我们可以在“基本信息”面板中修改采样频率和信号长度。一旦我们更改了任何数据，GUI将尝试生成合成PPG 并将其显示在对话框的底部。

通过按下“编辑”按钮，用户可以在弹出对话框中修改脉冲参数。对于常规PPG，此工具箱对不同脉冲使用相同参数。但对于不规则PPG，参数在早搏组的第一次搏动、早搏组的第二次搏动和参考搏动中是不同的。用户还可以修改第一心跳持续时间和第二心跳持续时间与参考心跳持续时间的比率。默认值如表1所示。编辑脉冲参数后，按“OK”按钮保存这些参数并进入主对话框。

在设置基本信息后，用户需要设置一些参数来生成脉冲持续时间。对于常规PPG，用户可以在“脉冲持续时间信息”面板中修改RR间期的平均心率和标准差。对于不规则PPG 类型，此面板更改为“不规则持续时间信息”面板，其中用户可以修改基本心率和不规则时间，

the synthetic PPG. The basic heart rate and mean heart rate are in the range of 50 to 180. A warning dialogue pops up when the "Irregular Times" value is too large or too small relative the signal length. In this case, users should either decrease the irregular times or increase the length of the signals.

If necessary, users can add noise to synthetic PPG. Two types of noise are available in the "Noise Info" panel: White Gaussian noise and multi-frequency noise. For white Gaussian noise, users can modify the signal-to-noise ratio (SNR). A 5-s regular PPG with white Gaussian noise is shown in **Figure 4A**. Additionally, for multi-frequency noise, see as **Figures 4B–D**. "Amplitude" is the amplitude of the peak of the noise signals, and "Frequency" describes the noise frequencies. The number of values in "Amplitude" and "Frequency" should be the same.

After synthesizing signals, users can press the "Save" button to save the synthetic PPG to comma-separated values file (.csv), Microsoft Excel file (.xlsx), and MAT-file (.mat).

6. LIMITATIONS OF STUDY AND FUTURE WORK

Generating regular PPG signal using certain parameters is reproducible. If we add noise, the PPG signal cannot be reproduced as the noise addition is carried out randomly. On the other hand, generating irregular PPG signals is non-reproducible because the duration of each beat is randomly getting set. Adding noise to the generated irregular PPG signal makes it highly non-reproducible.

The next step is to generate re-entry irregular heartbeats in PPG signals, and potentially other types of abnormalities to the toolbox. The main focus of the current study was not on detecting events in irregular PPG signals with irregular heartbeats; rather, the focus was on generating irregularity in PPG signals. Another aspect of future development is to generate PPG signals with certain hemodynamic parameters (e.g., blood pressure levels) simulating the PPG templates and their associated hemodynamics parameters. This toolbox is released as version 1 (PPGSynth v1.0, August 11, 2020) and the more templates we include the more the toolbox will be more able to generate PPG waveforms covering different irregularities (simulated cardiovascular patient groups) and noise types. One of the next steps is to generate normotensive and hypertensive PPG signals.

7. SUMMARY

PPGSynth, a new publicly available toolbox, is described as a means to generate synthetic PPG waveforms. Users can easily generate a waveform across a range of sampling frequencies and can also set the length of regular and irregular PPGs. The utility can also generate specific shapes of PPGs by modifying the pulse parameter settings. These characteristics make the new toolbox useful for less experienced users that would like to generate synthetic PPGs for their research and training in physiological measurements.

DATA AVAILABILITY STATEMENT

The PPGSynth MATLAB toolbox is publicly available. The code can be downloaded via this link: https://github.com/Elgendi/PPG-Synthesis/tree/master/code and the .exe file can be downloaded via this link: https://github.com/Elgendi/PPG-Synthesis/tree/master/exe.

AUTHOR CONTRIBUTIONS

ME designed the study, led the investigation, drafted the manuscript for submission with revisions, and feedback from the contributing authors. QT, ZC, JA, AA, CM, RW, and ME conceived the study, provided directions, feedback, and revised the manuscript. All authors approved the final manuscript.

FUNDING

This research was supported by the NSERC grant RGPIN-2014-04462 and Canada Research Chairs (CRC) program. RW was supported by NPRP grant # NPRP12S-0305-190231 from the Qatar National Research Fund (a member of Qatar Foundation). The findings achieved herein are solely the responsibility of the authors.

REFERENCES

1. Allen J. Photoplethysmography and its application in clinical physiological measurement. *Physiol Meas.* (2007) 28:R1–39. doi: 10.1088/0967-3334/28/3/R01
2. Elgendi M. On the analysis of fingertip photoplethysmogram signals. *Curr Cardiol Rev.* (2012) 8:14–25. doi: 10.2174/157340312801215782
3. Johnson AE, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. *Sci Data.* (2016) 3:160035. doi: 10.1038/sdata.2016.35
4. Liu D, Gorges M, Jenkins SA. University of Queensland vital signs dataset: development of an accessible repository of anesthesia patient monitoring data for research. *Anesthesia Analgesia.* (2012) 114:584–9. doi: 10.1213/ANE.0b013e318241f7c0
5. Charlton PH, Bonnici T, Tarassenko L, Clifton DA, Beale R, Watkinson PJ. An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram. *Physiol Meas.* (2016) 37:610–26. doi: 10.1088/0967-3334/37/4/610
6. Liang Y, Chen Z, Liu G, Elgendi M. A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in China. *Sci Data.* (2018) 5:180020. doi: 10.1038/sdata.2018.20
7. Roskamm H, Csapo G. *Disorders of Cardiac Function.* New York, NY: M. Dekker (1982).
8. Elgendi M, Eskofier B, and Abbott D. Fast T wave detection calibrated by clinical knowledge with annotation of P and T waves. *Sensors.* (2015) 15:17693–714. doi: 10.3390/s150717693
9. Tang Q, Chen Z, Ward R, Elgendi M. Synthetic photoplethysmogram generation using two Gaussian functions. *Sci Rep.* (2020) 10:13883. doi: 10.1038/s41598-020-69076-x

合成PPG 基本心率和平均心率在50 到180 之间。当"不规则时间"值相对于信号长度过大或过小时，会弹出警告对话框。在这种情况下，用户应该减少不规则的时间或增加信号的长度。

如有必要，用户可以向合成PPG 添加噪声。"噪声信息"面板中提供两种类型的噪声：白色高斯噪声和多频噪声。对于白色高斯噪声，用户可以修改信噪比（SNR）。图4A 中示出了具有白色高斯噪声的5 - s 规则PPG 。此外，对于多频噪声，参见图4B - D。"振幅"是噪声信号峰值的振幅，"频率"描述噪声频率。"振幅"和"频率"中的值数量应相同。

合成信号后，用户可按下"保存"按钮，将合成PPG 保存为逗号分隔值文件（.csv）、Microsoft Excel 文件（.xlsx）和MAT文件（.mat）。

6.研究和今后工作的局限性

使用某些参数生成规则PPG 信号是可再现的。如果我们添加噪声，则PPG 信号不能被再现，因为随机地执行噪声添加。另一方面，生成不规则PPG 信号是不可再现的，因为每个搏动的持续时间是随机设置的。将噪声添加到所生成的不规则PPG 信号使其高度不可再现。

下一步是在PPG 信号中生成再入不规则心跳，并可能将其他类型的异常生成到工具箱中。当前研究的主要重点不是检测具有不规则心跳的不规则PPG 信号中的事件;而是关注PPG 信号中的不规则性。未来发展的另一方面是生成具有某些血液动力学参数（例如，血压水平），其模拟PPG 模板及其相关联的血液动力学参数。此工具箱作为版本1（PPGSynth v1.0，8 月11日）发布2020），我们包含的模板越多，工具箱就越能够生成涵盖不同不规则性（模拟心血管患者组）和噪声类型的PPG 波形。接下来的步骤之一是生成血压正常和高血压PPG 信号。

7.总结

PPGSynth 是一个新的公开可用的工具箱，被描述为一种生成合成PPG 波形的方法。用户可以轻松生成一系列采样频率的波形，还可以设置规则和不规则PPG 的长度。该实用程序还可以通过修改脉冲参数设置生成特定形状的PPG 。这些特征使得新工具箱对于想要生成合成PPG 用于其生理测量研究和培训的经验较少的用户非常有用。

数据可用性声明

PPGSynth MATLAB工具箱是公开的。代码可通过以下链接下载：https://github.com/Elgendi/PPG-Synthesis/tree/master/code，.exe 文件可通过以下链接下载：https://github.com/Elgendi/PPGSynthesis/tree/master/exe。

作者贡献

ME设计了这项研究，领导了调查，起草了手稿，并提交了修订版，以及来自投稿作者的反馈。QT、ZC、JA、AA、CM、RW和ME构思了这项研究，提供了指导、反馈并修改了手稿。所有作者都同意了最终手稿。

资金

这项研究得到了NSERC 赠款RGPIN- 201404462 和加拿大研究主席（CRC）计划的支持。RW得到了卡塔尔国家研究基金（卡塔尔基金会成员）的NPRP 资助# NPRP 12S - 0305 - 190231 。本文所获得的结果完全由作者负责。

引用

1. 艾伦J.光电体积描记法及其在临床生理测量中的应用。生理学测量（2007）28：R1- 39 。doi：10.1088/0967 - 3334/28/3/R01

2. 永远的M。On the analysis of fingertip photoplethysmogram signals.（分析指纹摄影信号）Curr Cardiol Rev.（2012）8：14—25 。doi：10.2174/157340312801 215782

3. Johnson AE，Pollard TJ，Shen L，Lehman LH，Feng M，Ghassemi M，等人，MIMIC—III，一个免费的关键医疗数据库。SciData 。（2016）3：160035 。doi：10.1038/sdata .2016.35 ；

4. Liu D，Gorges M，Jenkins SA.昆士兰大学生命体征数据集：开发一个可访问的麻醉患者监测数据库用于研究。麻醉和镇痛。（2012）114：584 - 9。doi：10.1213/ANE.0b013e318241f7c0

5. 查尔顿PH，Bonnici T，Tarassenko L，克利夫顿DA，Beale R，Watkinson PJ.从心电图和光电容积描记图估计呼吸率的算法评估。生理学测量（2016）37：610 - 26 。doi：10.1088/0967 - 3334/37/4/610

6. 梁毅，陈志，刘刚，Elgendi M.一个新的，短记录的光体积描记图数据集，用于中国的血压监测。Sci Data.（2018）5：180020 .doi：10.1038/sdata .2018.20

7. Roskamm H，Csapo G.心脏功能障碍。纽约州纽约：M. Dekker（1982 年）。

8. Elgendi M、Eskofier B和Abbott D.通过临床知识校准的快速T波检测，带有P波和T波注释。传感器.（2015）15：17693 - 714。doi：10.3390/s150717693

9. 陈志华，陈志华，陈志华.使用两个高斯函数的合成光电体积描记图生成。（2020）10：13883 .doi：10.1038/s41598 - 020 - 69076 - x

10. Byrd RH, Hribar ME, Nocedal J. An interior point algorithm for large-scale nonlinear programming. *SIAM J Optimizat.* (1999) 9:877–900. doi: 10.1137/S1052623497325107

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

10. Byrd RH，Hribar ME，Nocedal J .大规模非线性规划的内点算法。SIAM J Optimizat .等（1999 ）9：877 - 900 。doi：10.1137/ S1052623497325107

利益冲突：作者声明，本研究是在不存在任何可能被视为潜在利益冲突的商业或财务关系的情况下进行的。