Draft: Unveiling the ActiLife Algorithm: Converting Raw Acceleration Data to Activity Count

Dr. Zhibiao Rao
Qmedic Health. Inc
4 School Street, Suite 350
Boston, MA 02108, USA
zrao@qmedichealth.com

Dr. Fahd Albinali
Qmedic Health. Inc
4 School Street, Suite 350
Boston, MA 02108, USA
falbinali@qmedichealth.com

ABSTRACT
All physical activity assessments are based on activity counts. The process of converting raw acceleration data to activity counts is lack of transparency and still commercial proprietary. The main purpose of this paper is to unveil this process and improve its transparency in the ActiLife software which is widely used in academic studies. The method to unveil the algorithm in the ActiLife software includes three aspects: filter shape, amplitude threshold and weighting factor. This paper identified these three aspects by investigating activity counts from the ActiLife which is fed with specific artificial data. Results show that the activity counts from the ActiLife and the unveiled algorithm match well with each other based on ten 20-days raw acceleration data. It concludes that the unveiled algorithm can be used to replicate the activity counts from the ActiLife.

Keywords
Accelerometer; activity count; ActiGraph; data filtration; physical activity.

1. INTRODUCTION
The term, activity count, is widely used in academic studies and industry. Activity counts, originally derived from the raw acceleration, are the foundation of physical activity assessments, such as wear and non-wear events identification (Choi et al, 2011). Ironically, very few people know what the activity count means and how it is computed from raw acceleration data. The signal processing techniques is obscure and owned by commercial software, such as ActiLife. The high-level signal processing procedures are: (1) the raw acceleration is firstly filtered by a bandpass filter; (2) the threshold is applied to set the value below it be zero; (3) A weighting factor is applied to scale the post-filtered data which is summed in 1 minute to get the activity count.

Some scholars tried to explore part of this process. Peach et al (2014) explored this process by generating hundreds of artificial sinusoidal acceleration data sampled at 30Hz and then input these signal into the ActiLife software to determine the nature of the ActiLife filtering algorithm. Brønd and Arvidsson (2015) determined the nature of the ActiLife filtering algorithm for data sampled at other frequencies. However, no public information is accessible for the whole process, from raw data to the activity count. The lack of transparency poses a challenge for comparison of historical accelerometer data in counts and data collected using raw accelerometer in S.I. units-m/s2 (Peach et al, 2014).

The main objective of this paper is to unveil the ActiLife algorithm which converts the raw acceleration data to the activity counts. The whole converting process mainly includes three stages: (1) Identifying the filter shape in the ActiLife by feeding hundreds of pure sinusoidal data with changing frequency into the ActiLife; (2) Identifying the amplitude threshold by feeding hundreds of pure sinusoidal data with changing amplitude into ActiLife; (3) Identifying the amplitude weighting factor by comparing activity count from the ActiLife with that from the unveiled algorithm in this paper.

2. ACTIGRAPH GT3X+
One activity monitor widely used in physical activity research is the ActiGraph GT3X+ which collects raw acceleration data. The measured raw acceleration is in a dynamic range of ±6*g* with sensitivity of 3*mg*/LSB (GT3X+ and wGT3X+ Device Manual, 2015). The default sampling frequency is 30Hz. The ActiLife 6 software, visualizing and processing acceleration data, reports two fundamentally different data types: the raw data reported as multiples of *g*, and counts, a proprietary and dimensionless unit (Peach et al, 2014).

3. METHOD TO UNVEIL THE ACTIFLIFE ALGORITHM
ActiLife is ActiGraph's premier actigraphy data analysis software platform for physical activity assessment. From the ActiGraph support website, activity counts are defined as a result of summing up post-filtered accelerometer values into epoch "chunks". The value of the counts will vary based on the frequency and intensity of the raw acceleration. The procedure of calculating activity counts from raw acceleration data in the ActiLife we guess is as below:

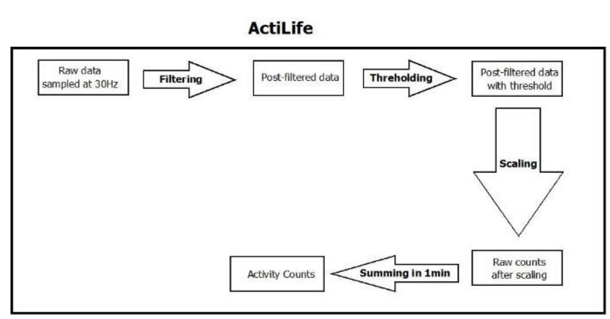

Figure 1. Schematic diagram of converting raw acceleration data to activity counts in the ActiLife software

The ActiLife algorithm acts in four stages. Firstly, a frequency bandpass filter is implemented to eliminate unwanted signals. Secondly, a preset threshold is applied to filter the value below it. Thirdly, a weighting factor is used to scale up the post-filtered acceleration data. The last stage is pretty straightforward and the activity counts are obtained by summing up the scaled post-filtered acceleration in one minute. The main objective of this paper is to explore first three stages and identify the filter shape, threshold value and weighting factor used in the ActiLife software.

3.1 Filter Shape

Filter shape is a function of frequency with a maximum magnitude 1, which attenuates or eliminates signals in a certain frequency range. From the convolution theorem, it is able to determine the precise form of a filter when enough input and post-filtered signals are provided. However, in the ActiLife software, a known input signal passes through an unknown filter and is then processed with unknown methods in few other steps, it becomes very challenging to determine the precise shape of this filter based on the input signal and compressed output signal because other factors (like threshold and weighting factor) may affect the filter shape.

When pure sinusoidal signals with a fixed amplitude but changing frequencies pass through the ActiLife software, the filter shape is only the function of frequency. Threshold, weighting factor and summing up method upon the post-filtered acceleration data scale down or up count per minute with the same weight, but the filter shape doesn't change. In a word, normalized activity count from pure sinusoidal signals are only affected at the filtering stage but not others, such as thresholding, scaling or summing up.

Pure sinusoidal signals with 1g amplitude at 150 specific frequencies are generated as an input to the ActiLife software:

$$Xacc = sin(2\pi f t) \quad (1)$$

1g amplitude is chosen here because the threshold is supposed to be smaller than 1g. Sampling frequency is set to be 30Hz which is the same as the default sampling frequency in the GT3X+ device. f is the response frequency which is from 0 to 5Hz with increment 1/30, and t is the time with 30 minutes long.

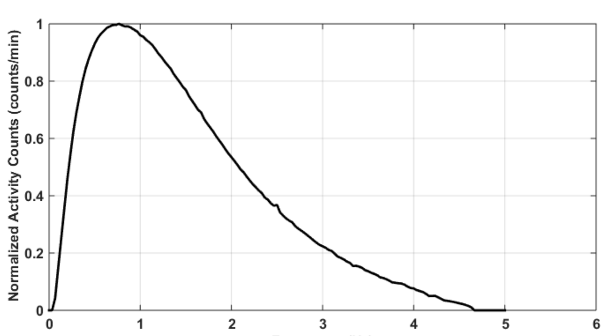

Figure 2. Activity counts normalized by the activity count at the peak frequency 0.77Hz. 50% cut-offs: 0.22-2.07Hz. 0% cut-offs: 0.05-4.67Hz.

The 150 artificial datasets are generated into .csv files which have the same format with GT3X+ raw data files. The count per minute can be read from 1-min epoch .csv file in the ActiLife software.

Figure 2 shows the activity counts normalized by the activity count at the peak frequency 0.77Hz. The activity counts are the output counts from the ActiLife software based on the raw acceleration data in Eq.(1). Figure 2 also represents the filter shape implemented in the ActiLife software. This filter attenuates the signal at all frequencies except the peak frequency 0.77Hz. When the input signal is pure cosine, the filter shape in Fig.2 wouldn't change.

3.2 Amplitude Threshold

The amplitude threshold is applied to set the value below it to be zero. When the amplitude of the sinusoidal signal is below the preset threshold, the whole signal becomes zero value and the activity count will be zero. Based on this concept, the boundary of zero activity count is the amplitude threshold.

A bunch of pure sinusoidal signals with a fixed frequency but changing amplitudes are generated and input into the ActiLife software. Here the fixed frequency is chosen to be the peak frequency at which the signal will not be attenuated with the filter. The amplitude is chosen from 0 to 7g here. The plot of output counts versus amplitudes of the sinusoidal signal shows the amplitude threshold.

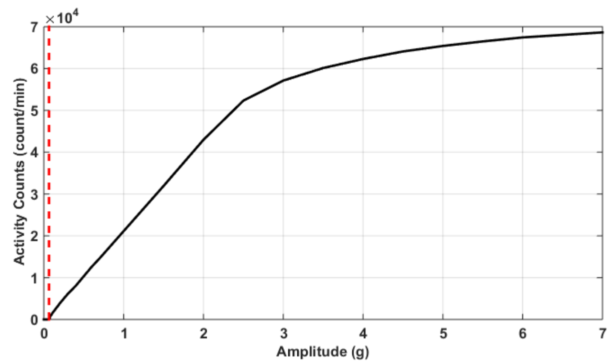

Figure 3. Activity count per minute from the ActiLife software versus the amplitude of sinusoidal input signals. The amplitude threshold is 0.065g. When the amplitude of the signal is below this threshold, the activity count is completely zero.

Figure 3 shows the activity count per minute changing with the amplitude. When the amplitude is below 0.065g, the activity count is completely zero which suggests amplitude threshold is 0.065g.

3.3 Weighting Factor

The raw signal, after being filtered with the filter identified in the section 3.1 and being applied with the threshold identified in the section 3.2, is converted to the raw count at each sampling point. The count per minute is the sum of absolute raw count in 1 minute. However this calculated count per minute doesn't match with the output activity count from the ActiLife software. A weighting factor may be applied. This weighting factor is different from the amplification factor which represents the conversion from an analog signal to a digital signal (Chen and Bassett, 2005).

Around 3-day measurement .gt3x data (320009309W.gt3x) is chosen as an example to identify the weighting factor. The principle to identify the weighting factor is to minimize the mean difference of activity counts from the ActiLife software and the unveiled algorithm. We denote the activity count from the ActiLife software by $AC_{ActiLife}$ and the unveiled algorithm in this paper by $AC_{Present}$ calculated by following the procedure in figure 6. The mean difference of count is defined as:

$$Mean\ difference\ of\ count = \left|Mean\left(AC_{ActiLife}\right) - Mean\left(AC_{Present}\right)\right| \quad (2)$$

Eq.(2) presents the calculation of mean difference of counts from the ActiLife software and from the unveiled algorithm. Its mean difference value versus the weighting factor is plotted in Figure 4.

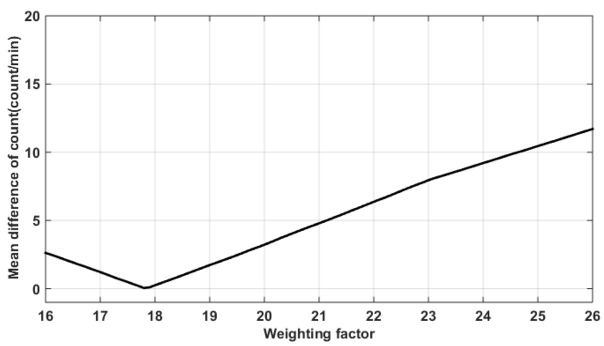

Figure 4. A plot of the mean difference of activity counts from the ActiLife software and from the unveiled algorithm versus the weighting factor. When the scale factor is 17.8, the mean difference value of activity count is minimum which suggests the weighting factor is 17.8.

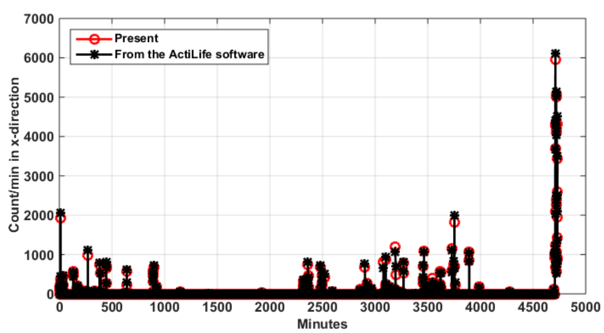

Figure 5. The comparison of activity counts from the ActiLife software and from the unveiled algorithm. The identified filter shape, threshold and weighting factor are implemented in the unveiled algorithm. Here only the activity count in the *x* direction is compared. Activity counts in other directions also match well.

Figure 4 shows the mean difference value of activity counts versus the weighting factor. It is clearly seen that the minimum mean difference value happens at the weighting factor 17.8. Once the filter shape, threshold and weighting factor of the ActiLife algorithm are identified, it enables to convert any raw acceleration data to the activity count. Figure 5 shows the comparison of activity counts from the ActiLife software and from the unveiled algorithm. These two activity counts are matching well.

The filter shape, threshold and weighting factor in the ActiLife software are systematically identified with both artificial and real data. Figure 6 shows the schematic diagram to replicate the activity count in the ActiLife software for any raw acceleration data.

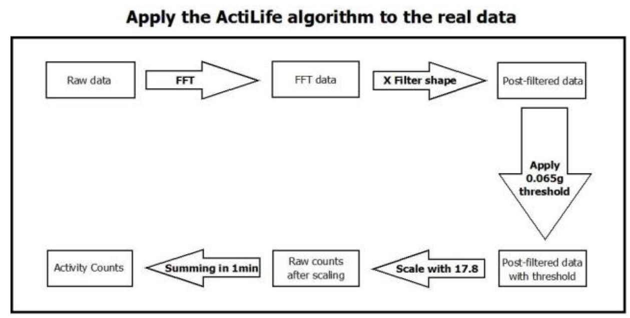

Figure 6. The schematic diagram of converting the raw acceleration data to the activity count with the unveiled ActiLife algorithm.

The main procedure to replicate the activity count with the unveiled ActiLife algorithm is as follows:

(1) Transfer the raw acceleration data from time domain to frequency domain by Fast Fourier Transfer (FFT)
(2) Multiply the filter shape identified in the section 3.1 to the FFT data
(3) Obtain the post-filtered time domain acceleration data by inverse FFT(IFFT) on the filtered data
(4) Set the value of post-filtered acceleration data to be zero when its absolute value is below the threshold value 0.065*g*
(5) Scale up the post-filtered acceleration data by multiplying the weighting factor 17.8.
(6) Sum up raw count in 1-min long window to get the activity counts

4. DISCUSSION

Figure 5 showed the activity count from the unveiled algorithm matches well with that from the ActiLife software. In this section, activity counts from raw acceleration data in three axes are investigated. Normally vector magnitude (VM) is the resulting vector that forms when combining the sampled acceleration from all three axes on any devices. The VM is defined as:

$$Vector\ Magnitude = VM = \sqrt{(AC_{Axis\ 1})^2 + (AC_{Axis\ 2})^2 + (AC_{Axis\ 3})^2} \quad (3)$$

Where $AC_{Axis\ 1}$, $AC_{Axis\ 2}$ and $AC_{Axis\ 3}$ represents the activity count in three axes.

Additional ten 20-day long datasets are processed to investigate the accuracy of the activity count from the unveiled algorithm compared to that from the ActiLife software. The results are shown in figure 7.

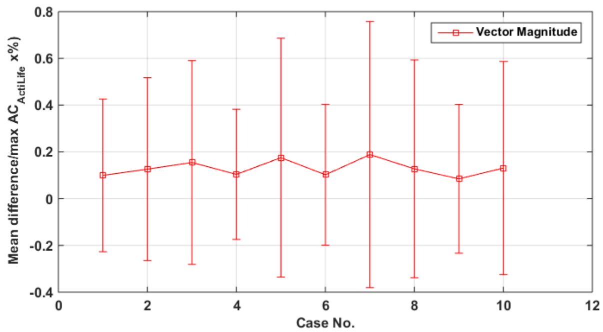

Figure 7. The ratio of mean difference to the maximum count from the ActiLife software in percentage. Here the mean difference value is computed from equation (2). The error bar represents its standard deviation around its mean value. The maximum error compared to the maximum activity count is below 0.8%.

Figure 7 shows the ratio of mean difference of count to the maximum count from the ActiLife software in percentage. The mean error is around 0.1%. For the GT3X+ device, its sensitivity is 3*mg*/LSB for the dynamic range of ±6*g* which means there exists quantization error. Quantization error results from trying to represent a continuous analog signal with discrete, stepped digital data. The quantization error appears in two stages: from the analog voltage signal to a digital signal and from the digital signal to signal in *g*. The total quantization error is 2x3mg/12g=0.05% which is comparable to the mean error 0.1% in figure 7.

The maximum error is below 0.8% which means the ActiLife algorithm has been explored well in this paper. The filter shape, threshold and weighting factors identified in this paper will help users to replicate the activity counts from raw acceleration data or develop their own algorithm. The unveiled procedure in figure 6 will help improve transparency in the ActiLife algorithm and help scholars understand how the activity count is calculated.

5. CONCLUSIONS

The term, activity count, is widely used in academic studies and industry. Activity counts, originally derived from the raw acceleration, are the foundation of physical activity assessments, such as wear and non-wear events identification. Ironically, very few people know what the activity count means and how to calculate from raw acceleration. The process of converting raw acceleration to activity count is still proprietary and kept as a secret by commercial companies, like ActiGraphy.

This paper has explored the ActiLife software and unveiled its process of converting raw acceleration to activity count step by step. By feeding hundreds of pure sinusoidal artificial acceleration data to the ActiLife software, the filter shape has been identified. Subsequently, the threshold has been found by feeding sinusoidal data with a fixed frequency but changing amplitude into the ActiLife software. At the end, the weighting factor is obtained by matching the activity counts from the unveiled algorithm with those from the ActiLife software. Additional ten 20-day raw acceleration datasets are selected to support that the unveiled algorithm works well in other large datasets. The small error indicates that the unveiled algorithm can be used to replicate the activity counts from the ActiLife.

With deep understanding in the process of converting raw acceleration to activity count, physiologists and public health researchers may develop their own filter or algorithms with more physical or physiological meaning.

6. ACKNOWLEDGMENTS

This work was financially supported by National Institute of Health.

7. REFERENCES

Brønd, Jan Christian, and Daniel Arvidsson."Sampling frequency affects ActiGraph activity counts." ICAMPAM Limerick Conference (2015).

Chen, Kong Y., and David R. Bassett. "The technology of accelerometry-based activity monitors: current and future." *Medicine and science in sports and exercise* 37.11 (2005): S490.

Choi, Leena, et al. "Validation of accelerometer wear and nonwear time classification algorithm." *Medicine and science in sports and exercise* 43.2 (2011): 357.

GT3X+ and wGT3X+ Device Manual. The ActiGraph, Pensacola; 2015.

Peach, D., J. Van Hoomissen, and H. L. Callender. "Exploring the ActiLife® filtration algorithm: converting raw acceleration data to counts." *Physiological measurement* 35.12 (2014): 2359.

Troiano, Richard P. "A timely meeting: objective measurement of physical activity." *Medicine and Science in Sports and Exercise* 37.11 (2005): S487.

Van Hees, Vincent T., Marcelo Pias, Salman Taherian, Ulf Ekelund, and Soren Brage. "A method to compare new and traditional accelerometry data in physical activity monitoring." In *World of Wireless Mobile and Multimedia Networks (WoWMoM), 2010 IEEE International Symposium on a*, pp. 1-6. IEEE, 2010.