A COLLECTIVE BIOLOGICAL PROCESSING ALGORITHM FOR EKG SIGNALS

Horia Mihail Teodorescu
FAS, Harvard University, Cambridge, MA, US
hmteodor@fas.harvard.edu

Keywords: swarm model, signal processing, filtering, biologic signal, EKG.

Abstract: We establish and explore an analogy between hunting by packs of agents and signal processing. We present a version of adaptive 'Hunting Swarm' algorithm (HSA), apply it to EKG signals, and investigate the influence of the model parameters on the filtering of stationary and nonstationary noise. We show that results obtained with the HSA filter may outperform results obtained with several other filters.

1 INTRODUCTION

Biological signals have wide bandwidth and may be affected by various noises (M. Beekman and Simpson, 2009). The first stage in processing such signals consists of filtering them in order to achieve a good signal to noise ratio (SNR). This task is often challenging because of the wide band of the signals and of the noises. As a consequence, numerous papers have been published recently proposing new filtering methods for ECG signals (Almenar and Albiol, 1999), (Leski and Henzel, 2005), (Kotas, 2007), (Korrek and Nizam, 2010), (Bansal et al., 2009), (Yan et al., 2010)

In this paper we explore several variants of the 'hunting swarm algorithm' (HSA) and analyze their ability to remove noise from EKG signals for various signal to noise ratios (SNR). The signal is 'enacted' by the trajectory of a prey hunted by the swarm, as detailed in section 2.

The models of the swarms in this paper include salient features from various swarm models reported in the literature and features that we introduced based on general considerations or from experimentation with model parameters.

The organization of the paper is as follows. In the second section we expose the method to transform the signal processing task into a pack-hunting-a-prey task and describe the equations describing the prey and the pack movements. The third section is devoted to the results of filtering ECG signals with the HSA algorithm. The details of the implementation and the results are discussed in the fourth section. Conclusions are drawn in the last section.

2 THE HS SIGNAL PROCESSING METHOD

2.1 Metaphor of the hunting pack

In this section we suggest and exploit an analogy between signal filtering and the natural hunting packs. We use this analogy to produce an algorithm for nonlinear signal processing. The analogy has two main players: the prey and the hunting pack. The prey does not collaborate to the signal processing; instead, it enacts the signal. The pack performs a virtual hunting and in so doing it produces the output (processed) signal as the trajectory of the center of the pack. The hunting pack model, while borrowing much from various swarm models, has many new features that give reason to consider it a new swarm model.

Figure 1 depicts a sketch of a simplified processing procedure. In this sketch, the swarm is assumed constrained on a line at each time moment, with the agents taking positions along that vertical line, according to movement equations governed by inter-agent forces and to agent to prey forces. The prey moves in discrete time along the signal. The agents

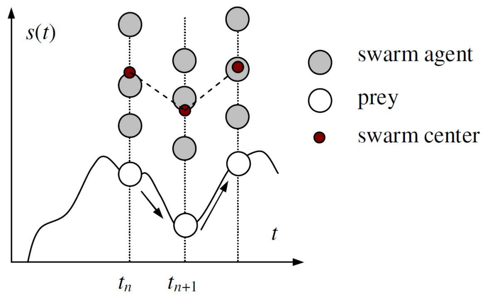

Figure 1: Sketch of the operation of the swarm as signal processing system.

are attracted by the prey, thus tending to follow the prey. Consequently, the center of the swarm describes a trajectory in the plane. That trajectory is the result of 'processing' the prey trajectory, i.e. the signal, by the swarm.

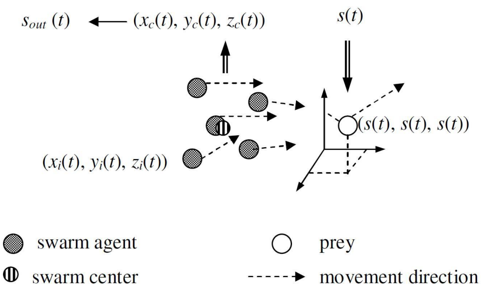

Figure 2: Sketch of the operation of the swarm as signal processing system.

Figure 2 shows a sketch of the procedure in three dimensional (3D) space. In this sketch, the prey moves along a trajectory represented by $(x_p(t) = s(t), y_p(t) = s(t), z_p(t) = s(t))$ and represents the input signal marked by the double line arrow in the upper right side of the figure, while the center of the pack represents the output signal.

The use of this metaphor in signal de-noising is based on the hypothesis that hunting swarms are able to filter out 'undue', 'evasive', that is, noise-like changes in the trajectory of the prey during the hunting. Moreover, swarms might use a simple collective adaptation of its behavior to closely follow the prey when the last had the chance to take a larger distance. These hypotheses were verified during simulations, as demonstrated by the result section. The consequence is the present proposal of HS filtering method.

Because the model is somewhat elaborate, we introduce in the subsequent section the equations of the swarm, neglecting the prey, while in the 2.3 subsection we take into account the prey influence and the adaptive behavior of the prey as elicited by the prey movement.

2.2 Swarm basic equations

Hunting takes place according to a set of equations that govern the movements of the agents in the pack. These equations have three types of components. The first type comprises 'physical' forces like the inertia and the friction forces. The second class of forces includes the interaction forces inside the pack; these forces keep the pack together, while preventing agents from colliding one with the other. The agents are endowed with an elementary memory and with an awareness to the global state of the pack, moving accordingly. Finally, the 'external' force that produces the movement of the swarm is the interaction of the agents with the hunted prey. The general equation governing a swarm is, according to (Reza Olfati-Saber and Murray, 2007):

$$\dot{x}_i(t) = \sum_{j \in N_i} (x_j(t) - x_i(t)) + b_i(t) \tag{1}$$

with the initial conditions $x_i(0) = z_i$ and $b_i(t) = 0$. Above, x is a spatial coordinate, i, j denote agents of the swarm, b is due to an external force (bias). The left hand side of the equation represents the velocity, while the terms under the sum in the right-hand side are similar to elastic forces, $F = \kappa \cdot (x_0 - x)$, where κ is the elastic constant and x_0 is a fixed position. For the swarm, x stands for the position of the agent x_i and x_0 is replaced by that of the prey, x_p. In case of consensus algorithms, the above equation in discrete time and without the term b_i is (Reza Olfati-Saber and Murray, 2007):

$$x_i[t+1] = x_i[t] + ct \cdot \sum_{j \in N_i} a_{ij} \cdot (x_j[t] - x_i[t]) \tag{2}$$

where a_{ij} are constants and t denotes here a discrete time moment. Under certain conditions, the swarm is stable, which in terms of consensus theory means that a consensus is asymptotically reached (Reza Olfati-Saber and Murray, 2007).

Kim (Kim, 2008) used artificial potential functions to model the attraction towards the goal. We use a similar approach, but with different potential functions, moreover also including repulsive forces that replace the attractive ones starting with a given distance. The potential forces we use have the form:
(i) - for the repulsive forces:

$$F_{i,j} = -k_1 \cdot \frac{x_j - x_i}{d_{i,j}^{\eta_1}} \tag{3}$$

for $d_{i,j} \leq \rho_1$, where k_1 is a positive constant, $d_{i,j}$ is the distance between the agents denoted by the indices i and j, η_1 is a natural power, and ρ_1 is a constant;
(ii) - for the attractive forces:

$$F_{i,j} = k_2 \cdot \frac{x_j - x_i}{d_{i,j}^{\eta_2}} \quad (4)$$

for $d_{i,j} \leq \rho_1$, where k_2 is a positive constant, $d_{i,j}$ is the distance between the agents denoted by the indices i and j, η_2 is a natural power, and $\rho_2 > \rho_1$ is a constant. The constants in the above equations are parameters of the processing system.

The above forces create accelerations that are computed as:

$$a_{u,i}[t+1] = -k_1 \cdot \sum_j \frac{u_j[t-1] - u_i[t-1]}{d_{i,j}^{\eta_1}[t-1]} \quad (5)$$

where u stands for x, y, or z and $a_{u,i}$ is the acceleration in the direction u of the agent i due to the repulsion at small distances from the other agents in the pack, or due to attraction at larger distances. Subsequently, we use the first order approximations of the derivative, $\dot{u}[t] = (u[t] - u[t-1]) \cdot \delta$, where u is a coordinate variable, t is a discrete variable standing for time, and δ is the step for time discretization. Then, the inertial force along the u direction is $m \cdot \ddot{u} = m \cdot (v_u[t] - v_u[t-1])$, where v_u is the velocity along the u direction and m is the mass, which we assume unitary for all agents in the pack. Based on the acceleration, according to the last equation, the change of velocity is computed as

$$v_{u,i}[t+1] = v_{u,i}[t] + \delta \cdot a_{u,i}[t]. \quad (6)$$

In (6), δ is the time step interval and represents an important parameter in the simulations. Larger values of δ make the pack respond faster to the signal, but can produce overshoots when the signal varies fast. We used values of δ between 0.5 and 2 for best results.

Next, we include in equation 6 the effect of friction forces, that we assume to have components proportional to the respective velocity component, $F_{u,friction} = \mu \cdot v_u$. The change in velocity due to the friction is $\Delta v_u = F_{u,friction} \div m = -ct. \cdot v_u$, where the constant includes μ, the time step, δ, and the inverse of the mass, m^{-1}. For ease of writing, subsequently we denote by μ the constant in the change of the velocity due to friction, $\Delta v_u[t+1] = -\mu \cdot v_u[t]$.

2.3 Prey influence and adaptive swarms

The prey is assumed to move independently of the movement of the hunting swarm. This hypothesis is unsuitable for biological or physical modeling purposes, but it is required by the task we deal with, because the signal, enacted by the prey, should remain independent of the processing. On the other side, the prey 'attracts' the hunting swarm. The attraction force we use is a third order, nonlinear, elastic-type force with the expression:

$$F_{u;a,p} = A_1 \cdot (u_p - u_a) + A_2 \cdot (u_p - u_a)^3, A_1, A_2 \geq 0 \quad (7)$$

where u_p are the coordinates of the prey and A_1, A_2 are model constants. Including the contribution of the prey to the acceleration of the agents, the equation (6) rewrites

$$v_{u,i}[t+1] = v_{u,i}[t] + \quad (8)$$

$$+\delta \cdot a_{u,i}[t] + A_1 \cdot (u_p - u_a) + A_2 \cdot (u_p - u_a)^3. \quad (9)$$

where, again, we assume that we included the δ factor in the constants A_1, A_2 without changing the notations. The position of the agent at time step $t+1$ is obtained as

$$u_i[t+1] = u_i[t] + \delta \cdot v_{u,i}[t]. \quad (10)$$

A set of restrictions, like a limit in acceleration and a limit in the change of direction are added, which have intuitive biological counterparts. We skip details here, but we used these limits in the swarm processing system whose results we describe.

Once the positions of the N swarm agents at time t are computed, we determine the position of the center of the swarm as

$$u_s[t] = \frac{1}{N} \cdot \sum_i u_i[t]. \quad (11)$$

It is natural at the biological level that agents in the swarm are aware of the behavior of the swarm as a group and to adjust to it. We make a further hypothesis, that in a hunting pack the agents are aware of the relative position of the pack and the prey, adjusting their speed according to that relative position. Namely, we assume that, whenever the distance from the center of the pack to the prey becomes too large, every agent will increase its velocity by a factor proportional to $u_p - u_s$. So, if $|u_p - u_s| \geq D$, an increase in velocity $\Delta v_{u,i} = B \cdot (u_p - u_s)$ occurs for all the agents. This conditional increase of the agents velocity stands for an elementary adaptation to the momentary conditions of hunting. From the point of view of the HS filtering algorithm, this adaptive behavior means better results in case the signal has fast transients or fronts. We skip technical details related to the algorithm implementation and provide some of them in the Appendix.

3 PROCESSING RESULTS

We exemplify the results we obtained with the HS signal processing method applied to EKG signals. We used the signal database PhysioBank (Goldberger et al., e 13). All input signal filenames referred to in the figures are from Physiobank, to which we add noise corruption. We show two categories of results. The first one refers to noisy EKG from the cited database; the second refers to filtering results obtained when applying the processing to relatively clean EKG signals that we corrupted with uniform noise of various amplitudes.

3.1 Mechanics of the filtering process

The mechanics of the HS processing is revealed by the representation of the trajectories of all the agents in the pack and by a representation of the dependency of the evolution of the center of the pack with respect to the processed signal. The 'hunting' process has two phases. In the first phase, the pack, which is assumed to start from random initial conditions, is structuring itself and evolves toward an almost stable configuration. This transitory regime is shown in Fig. **??** and may last about 100 time steps, its duration primarily depending on the initial positions and on the friction forces. After the transitory behavior, the swarm remains almost stable, despite its continuous movement driven by the prey. Only when the signal has very fast variations, the swarm may be partly de-structured and needs some time to recover its equilibrium. This regime of dynamic stability is shown in Fig. **??** for a swarm including 55 agents.

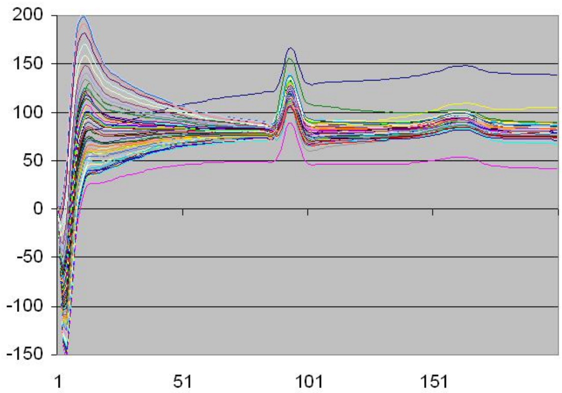

Figure 3: Transitory regime of the hunting pack takes about 100 time steps for this swarm of 55 agents, with $\mu = 0.35$, $\eta_1 = \eta_2 = 4$.

The representation of the trajectory of the center of the swarm as an implicit function of the trajectory of the prey shows, for almost all processed signals, that two or three regimes occur during the 'hunting', regimes that are represented by the loops in the diagram in Fig. **??**.

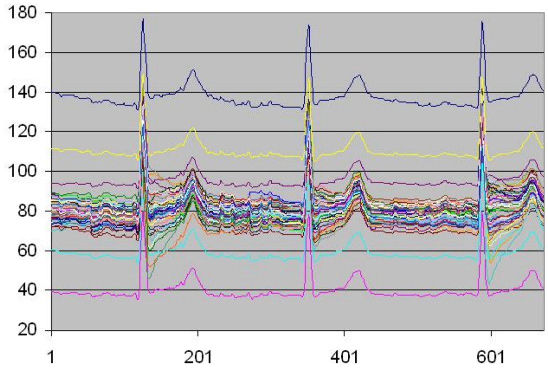

Figure 4: Processing result with a swarm with 55 agents, using the fourth power of distances in the inter-agent repulsive and attraction forces and a friction coefficient $\mu = 0.35$.

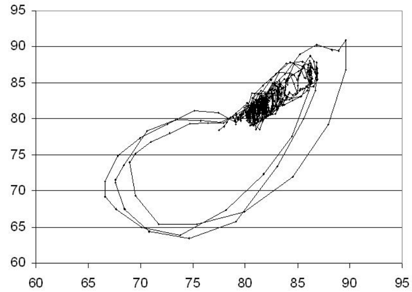

Figure 5: Swarm trajectory plot versus the signal.

3.2 Filtering noisy signals

For determining the usefulness of HS filtering, we tested the swarm filters with signals from the benchmark database PhysioBank ATM (Goldberger et al., e 13). Two types of tests were carried on: (i) filtering signals from PhysioBank that are noisy, and (ii) filtering clean signals to which controlled noise is added. The first type of tests is needed for determining if the new filters are able to solve a real-life problem; the second type of tests allows us to investigate the capabilities of the filtering procedure under various controlled conditions.

We exemplify the filtering of noisy signals from PhysioBank ATM with the signals Fantasia f1y07 and Apneea ECG A01.

The trajectories of the agents of a swarm of 55 agents during the 'hunting' process (Figure 4) are averaged to obtain the center of the pack trajectory. Results of the HS filtering are shown in Figures .

The various parameters of the swarm influence the results. For example, as expected, a too large friction coefficient would produce a slower 'catching' of the signal when the signal has large swings, like the QRS complex. However, the slowing down is not the same on the upper and lower front of the impulsive signal, because of the nonlinearity in the swarm behavior. This is seen in Fig. **??**.

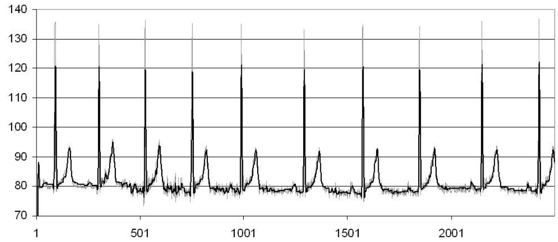

Figure 6: Processing result with a swarm with 55 agents, using the fourth power of distances in the inter-agent repulsive and attraction forces and a friction coefficient $\mu = 0.35$.

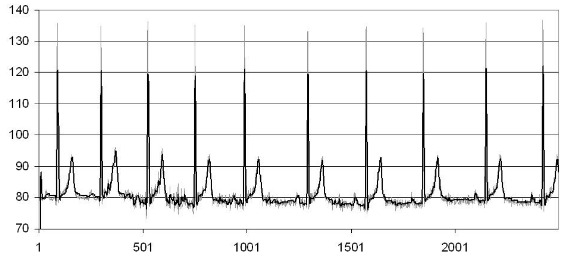

Figure 7: $\mu = 0.35$ Original signal and HS filtering result with a swarm with parameters described in the text.

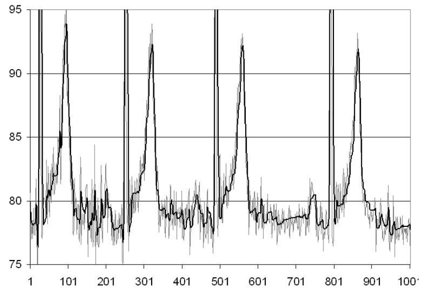

Figure 8: Detail from the above figure ($\mu = 0.65$).

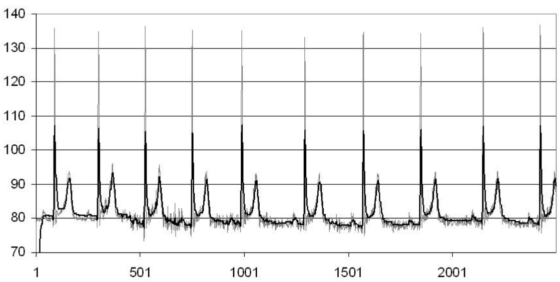

Figure 9: The same signal filtered with a swarm with friction coefficient $\mu = 0.65$.

The HS filters are causal, meaning that they take into account only previous values of the signal to generate the current value of the filtered signal. As a consequence, there is a delay between the produced output value and the current value of the signal. For the numerical evaluation of the performance of the HS filters, for example for applying the mean square error criterion, we need to determine the lag of the filter. We determined the corresponding lag for a filter with specified parameters minimizing the MSE between the signal and the HS output, for signals not corrupted with noise, according to the formula:

$$\tau = \min_{k}\left(\sum_{t=t_1}^{t_2} (s_c[t+k] - s_0[t])^2\right) \qquad (12)$$

where t_1, t_2 are the limits of the interval of determination (we used $t_1 = 200, t_2 = 950$), s_c is the output signal, and s_0 the input signal. The results for such a filter are given in Table 1, showing that the lag of this filter is $\tau = 3$. We removed the lag when we computed the MSE for the filtering of noisy ECGs.

Table 1: MSE errors of the swarm filter, for various adjustments of the lag. Signal Apnea ECG A01

delay	$\tau = 0$	$\tau = 3$	$\tau = 5$	$\tau = 6$
Total error	6247.8	2624.1	5971.8	7597.9
MSE	2.886	1.870	2.822	3.183

Table 2: MSE errors of the swarm filter (delay $\tau = 3$), average and median filters for the same signal and noise, for various adjustments of the lag. Signal Apnea ECG A01

Filter	HS $\tau = 3$	average	median
MSE	1.870	2.467	2.621

For quantitatively comparing the results of the HS filter with the results of median and average filters, we run the program over signals corrupted by us with uniform noise. The same input signal with the added noise was then filtered with an average filter and with a median filter, both of them with window width of length 11, centered on the current sample. When the noise is high, the HS filter outperforms in terms of MSE the other filters, as shown in Table 2.

4 DISCUSSION

The HS processing method is highly nonlinear and hence sensitive to the amplitude of the signals. Good results are obtained with the parameters we used for signal amplitudes in the ranges seen in the figures. We multiplied all signals in the cited database by a factor of 10 before processing.

The HS filters are causal and the processing results have a lag with respect to the signal. To determine the lag, we shifted the result and compared the sum of squared errors obtained for various shifts. The lowest error was obtained, in case of the signal Apnea-ECG A01 (length 10 seconds, data format: standard) for a lag of 3 time steps. The total squared error was computed for the various filters for 750 samples, for the samples from 200 to 950. We skipped the first 200 samples to avoid the transitory regime of the swarm filter. The mean square error, MSE, was determined as $MSE^2 = (\sum_{t=200..950}(s_c[t] - s_0[t])^2)/750$. The results related to the determination of the lag are shown in Table 1 in the Annex.

While the algorithm is $O(n)$ in the number of input signal samples, the calculations at each step involve looping over the swarm, moreover involve many multiplications. As a result, the processing is time consuming. A swarm of 55 agents, with $\eta_1 = 4$ and $eta_2 = 4$, implemented in a C++ unoptimized program that also writes more that 10 files on the disk, takes about 3 seconds to process 2500 samples of input signal. This means that the process can be performed in real time for ECG signals at a sampling frequency of about 800 Hz.

The HS filter produces smoother output than the average and median filters of order 11 (see Appendix).

The results are not exactly the same when the code is run several times. The method is not perfectly deterministic, as the swarm starts with random conditions, moreover several configurations of the swarm may have the same or similar internal energy, thus allowing the swarm to follow close but not identical trajectories when following the same prey.

The system is not guaranteed stable. For example, swarms with 25 agents or with 85 agents, the other parameters being the same as above, are unstable. As far as the swarm remains stable, the number of agents in the swarm was found to have less influence on the filtering error than parameters like μ and constants in adaptation.

While we used the analogy with the hunting process, the presented algorithm might be regarded as a social process of agreement of a group with a model, represented by the signal. While the analogy is similar with the one of swarms with leaders, it is still different, because the leaders are assumed to be influenced by the rest of the group, while the model acts independently from the behavior of the 'followers' group.

5 CONCLUSIONS

The HSA is essentially a new nonlinear filtering algorithm derived as a combination of several approaches in the literature. It might be a strong candidate in filtering signals with non-stationary, wide bandwidth noise.

The hunting swarm method may work remarkably well when the parameters of the swarm are trimmed according to the processed signal and noise peculiarities. The main advantage is that the HS filters leave the signals that have fast as well as slowly varying regions only slightly altered, while removing a consistent part of the noise. In this respect, we found that the HS filters behave better than the basic average and median filters.

ACKNOWLEDGEMENTS

I thank Dr. David Malan and Professor Leslie Valiant for essential advice and critics.

REFERENCES

Almenar, V. and Albiol, A. (1999). A new adaptive scheme for ecg enhancement. *Signal Processing*, 75(3):253 – 263.

Bansal, D., Khan, M., and Salhan, A. K. (2009). A computer based wireless system for online acquisition, monitoring and digital processing of ecg waveforms. *Computers in Biology and Medicine*, 39(4):361 – 367.

Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., and Stanley, H. E. (2000 (June 13)). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220. Circulation Electronic Pages: http://circ.ahajournals.org/cgi/content/full/101/23/e215.

Kim, D. H. (2008). Self-organization of swarm systems by association. *International Journal of Control, Automation, and Systems*, 6(2):253–262. April 2008.

Korrek, M. and Nizam, A. (2010). Clustering mit-bih arrhythmias with ant colony optimization using time domain and pca compressed wavelet coefficients. *Digital Signal Processing*, 20(4):1050 – 1060.

Kotas, M. (2007). Projective filtering of time-aligned ecg beats for repolarization duration measurement. *Computer Methods and Programs in Biomedicine*, 85(2):115 – 123.

Leski, J. M. and Henzel, N. (2005). Ecg baseline wander and powerline interference reduction using nonlinear filter bank. *Signal Processing*, 85(4):781 – 793.

M. Beekman, G. A. S. and Simpson, S. J. (2009). *The Book*. The publishing company, London, 2nd edition.

Reza Olfati-Saber, J. A. F. and Murray, R. M. (2007). Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233.

Yan, J., Lu, Y., Liu, J., Wu, X., and Xu, Y. (2010). Self-adaptive model-based ecg denoising using features extracted by mean shift algorithm. *Biomedical Signal Processing and Control*, 5(2):103 – 113.

APPENDIX

In the appendix we present a few more graphical results of filtering, with details.

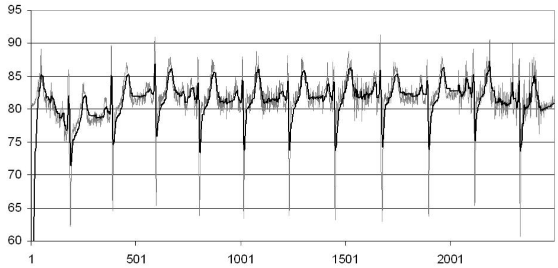

Figure 10: Result for the signal Fantasia f1o03 with a AHS with $\eta_1 = 3$, $\eta_2 = 4$, $\mu = 0.75, \eta_1 = 3, \eta_2 = 4$.

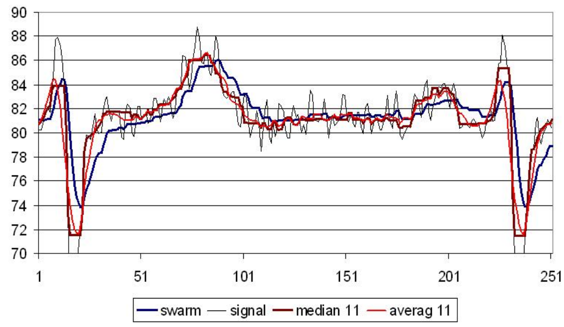

Figure 11: Comparison of median, average and swarm filters. The swarm produces a slightly smoother signal. Signal Fantasia f1o03, $\eta_1 = 3$, $\eta_2 = 4$, $\mu = 0.75$.

Comparison of swarm filter, average filter and median filter. The parameters of the swarm are: $Nmax = 55$, number of time steps $N_{Timesteps}$ = 1001, $\delta = 1.20$, $\gamma = 0.850, zeta = 3.89, phi_{max} = 0.785398$, friction coefficient $\mu = 0.250$, amplification inter-agent A = 2.40, amplification prey-agent AA = 0.80, amplification prey-agent second order term AB = 0.000020, amplification prey-agent AC = 0.60, amplification adaptive center swarm C = 1.30, adaptation distance D = 1.20, Noise amplitude Anoise = 0.0005.

The average, median and swarm filters have been applied with rectangular windows.

Table 3: MSE errors of the swarm filter (delay $\tau = 3$), for various values of the friction coefficient. Signal Apnea ECG A01

μ	0.15	0.20	0.25	0.30	0.35
MSE	2.213	1.944	1.870	1.935	2.009

The next figure shows details of AHS filtering the signal Fantasia 1o03

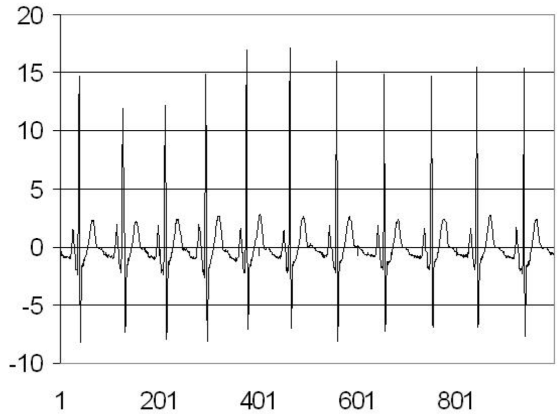

Figure 12: Comparison of median, average and swarm filters. The swarm produces a slightly smoother signal. Signal Apnea ECH A01 ($\mu = 0.25$, $\eta_1 = 3$, $\eta_2 = 4$, noise factor 0.0005.

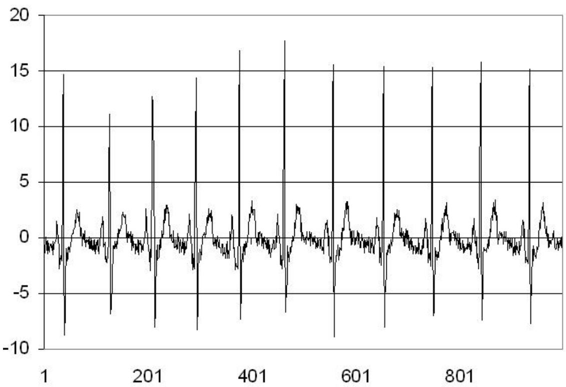

Figure 13: Noisy signal Apnea ECG A01, noise factor 00005.

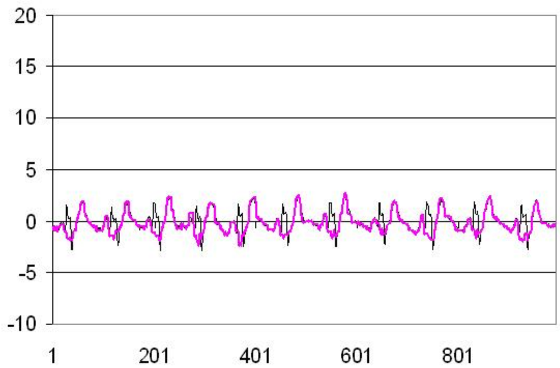

Figure 14: Average and median filters, window of 11 samples, centered. Signal Apnea ECG A01, noise factor 00005.

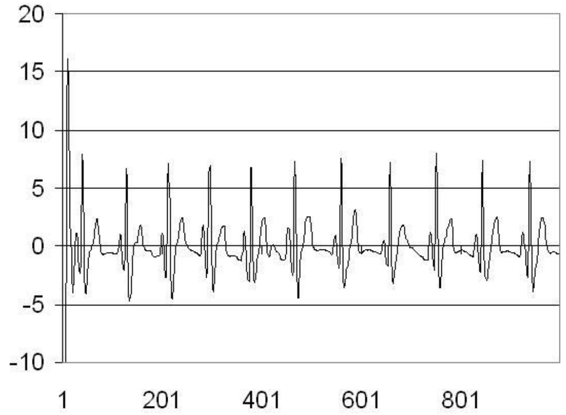

Figure 15: Result of filtering the signal Apnea ECG A01 with the parameters $\mu = 0.25$, $\eta_1 = 3$, $\eta_2 = 4$ for noise 0.0005.

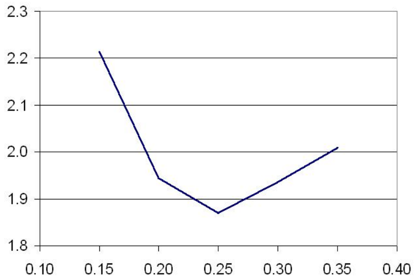

Figure 16: Variation of MSE as a function of the friction coefficient μ. Signal Apnea ECG A01

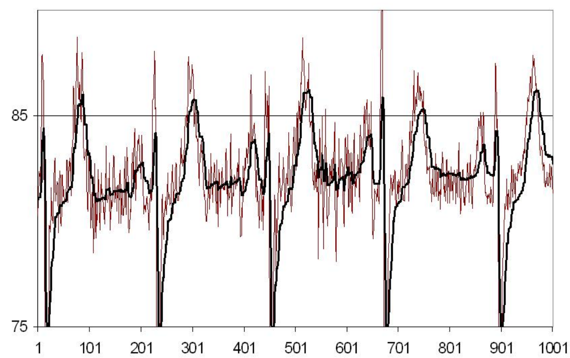

Figure 17: details of AHS filtering the signal Fantasia 1o03. Friction coefficient $\mu = 0.75$.