

# AI Search for Relaxing Music with Skin Conductance

Steven Lowes

5th October, 2018

## 1 Background

Music has the power to cause intense emotions in the listener [1], and the power to completely relax them [2]. Its power in manipulating emotions has led to widespread use in therapy, a field known as musical therapy [3].

The music used in music therapy is chosen specifically for the individual patient, respecting the fact that listeners have different responses to the same music [3]. Research into song selection for music therapy is based on manual extrapolation from listeners self-reporting which elements of the music they find relaxing.

We aim to replace the need for a music therapy professional by implementing a system that dynamically searches for songs that an individual finds relaxing. To measure how relaxed the listener is, we can use the Galvanic Skin Response, also known as Electrodermal Activity.

## 2 Proposed Solution

Using the Spotify API, we can represent a song as a 14-dimensional vector which describes its important qualities, including *danceability*, *energy*, and *key* [4, 5]. We have a database of approximately 500,000 of the most popular songs on Spotify, and the ability to search them based on their vector representations.

We model the problem as a Global Optimization Problem. The set of solutions are the 500,000 songs in the database. To determine the fitness of any solution, we play it to a listener, and measure the change in Skin Conductance. A good solution should lower the listener's skin conductance.

## 2.1 Assumptions

The following assumptions have been made and will need to be tested and confirmed before the solution can be implemented:

- The fitness of one solution is roughly the same when it is tested twice (there is little error in the fitness function)
- Solutions which have similar vectors usually have similar fitness
- A song with high fitness correlates with the listener self-reporting as feeling relaxed

## 2.2 Areas of Concern

- Testing the fitness of one solution will require listening to an entire song. This means each invocation of the fitness function will take approximately 3 minutes.
- This results in an upper bound of 100 invocations of the fitness function per search, equivalent to 5 hours of music listening.
- Since we want to demonstrate differences in preference between listeners, we cannot run the fitness function in parallel with multiple listeners.

There are some mitigating factors, which reduce the chance that these areas of concern will prevent the solution from working:

- We know that two solutions which are similar have similar fitness, so we can reduce the number of invocations required by performing clustering analysis.
- We have essentially unlimited computation power for anything that does not involve a human listener. Even complex solution analysis will run quickly with only 500,000 solutions to test.
- We may be able to play only some of a song to test its fitness. This would speed up execution. The Spotify API can provide 30-second previews of songs [6], which usually align with the chorus of the song.

## 3 Literature Survey

### 3.1 Music as Vectors

Spotify API allows us to convert songs to vectors [4]. Echo nest has more documentation on this [5].

Other work that has done clustering analysis has found that the clusters mostly match up with what we expect (genres), so that suggests that one of our assumptions (similar vectors = similar songs = similar fitness) is true [7, 8].

### 3.2 Skin Conductance

Skin Conductance will be measured via a Grove GSR Sensor [9], which can be read via an arduino into a PC. The GSR sensor measures the activity in the sympathetic nervous system. When a person is relaxed, there is very little activity in the sympathetic nervous system, and when they are experiencing intense emotion, positive or negative, the sympathetic nervous system is very stressed and active, which causes subconscious sweating. We can measure this by testing the skin's conductance. When the skin is highly conductive, we know the listener is sweating more, and therefore experiencing more intense emotion [10]. The same sensors are used in detection of stress by polygraph systems [11].

### 3.3 Fitness Landscape

Fitness Landscapes are a method of visualising the fitness of a solution, which can be used when there is some notion of dimensional nearness of two solutions [12]. In our case, we can use the vector values of a solution as its dimensional position. A working solution to our problem will also produce a system which can output a fitness landscape for an individual, showing which songs they find relaxing and which songs they find stressful. These fitness landscapes could be used for further research, and comparison between individuals.

### 3.4 Global Optimization Search

Our fitness function consists of playing a song to a listener and measuring their change in Skin Conductance. This means that each invocation of the fitness function takes on the order of minutes to complete. The fitness function cannot be calculated in parallel, due to the need for a human listener, and we cannot use multiple human listeners due to the differences in individual preference that we are testing.

Multiple state-of-the-art AI Search techniques were considered, but all were deemed to be too time-complex for our use case. These included Tabu Search, Hill Climbing, Simulated Annealing, Genetic Algorithms, Newton's Algorithm [13], and the Bee Algorithm [14].

It is clear that we need to reduce the time complexity of our search, even at the cost of a very non-optimal solution. A StackExchange user lead me to research into the solving of *Expensive Black-Box Optimisation Problems* (EBOPs) [15].

These EBOP-solving algorithms use methods such as *clustering* and *surrogates* to reduce the number of times we need to invoke the expensive fitness function [16, 17, 18].

### 3.5 Clustering

Clustering can be used to reduce the number of function invocations needed in our search [16, 19]. Clustering our solutions means that we can reduce the number of solutions without reducing the amount of diversity in the set of solutions.

Clustering is computationally resource-intensive. However, we are not concerned about reducing the amount of computation done, as the time cost of listening to songs (fitness function invocations) is far higher than the time cost of any clustering algorithm will be.

### 3.6 Bayesian & Surrogate Techniques

These techniques work by constructing a model of what we expect the fitness landscape to be, and iteratively updating the predicted landscape each time more data becomes available, after a fitness function invocation [20, 18]. By maintaining a model of the predicted fitness landscape, we can test solutions that are predicted to be the best. This allows us to stop our search early, once we have found that the predicted fitness is often matching the actual fitness. This saves us from having to do an exhaustive search, while maintaining that the best solution found is probably the best solution globally [16].

These techniques are computationally expensive, but we are not concerned with computational costs due to our relatively small set of solutions. The time taken for computation of predicted fitness landscapes should never outweigh the time taken for the subject to listen to the music.

## References

- [1] P. V. Agostino, G. Peryer, and W. H. Meck, “How music fills our emotions and helps us keep time,” *Behavioral and Brain Sciences*, vol. 31, pp. 575–576, Oct. 2008.
- [2] D. Elliott, R. Polman, and R. McGregor, “Relaxing Music for Anxiety Control,” *Journal of Music Therapy*, vol. 48, pp. 264–288, Oct. 2011.
- [3] American Music Therapy Association, “Definition and Quotes about Music Therapy | Definition and Quotes about Music Therapy | American Music Therapy Association (AMTA).”

- [4] Spotify, “Get Audio Features for Several Tracks | Spotify for Developers.”
- [5] T. Jehan and D. DesRoches, “Analyzer Documentation.”
- [6] Spotify, “Get a Track | Spotify for Developers.”
- [7] J. D. D. Santos, “Discovering similarities across my Spotify music using data, clustering and visualization,” Oct. 2017.
- [8] “CoachellaR - A Cluster Analysis | RCharlie Analytics.”
- [9] SeeedStudio, “Grove - GSR sensor - Sensor - Seeed Studio.”
- [10] B. Farnsworth, “What is GSR (galvanic skin response) and how does it work?,” July 2018.
- [11] D. M. TUCKETT, “Detection of Deception in the Laboratory as a Function of Motivation and Set Size (polygraph Tests, Skin Conductance, Lie Detector, Guilty Knowledge Test),” Master’s thesis, Western Michigan University, United States – Michigan, 1986.
- [12] C. M. Reidys and P. F. Stadler, “Combinatorial Landscapes,” *SIAM Review*, vol. 44, pp. 3–54, Jan. 2002.
- [13] A. Chandel and M. Sood, “Searching and Optimization Techniques in Artificial Intelligence: A Comparative Study & Complexity Analysis,” vol. 3, no. 3, p. 6, 2014.
- [14] D. T. Pham, A. Ghanbarzadeh, E. Koc, S. Otri, S. Rahim, and M. Zaidi, “Bee Algorithm A Novel Approach to Function Optimisation,” p. 41.
- [15] StackExchange, “search algorithms - Optimization Problem when calculating fitness is expensive,” Oct. 2018.
- [16] F. Boukouvala and M. G. Ierapetritou, “Derivative-free optimization for expensive constrained problems using a novel expected improvement objective function,” *AIChE Journal*, vol. 60, pp. 2462–2474, July 2014.
- [17] D. R. Jones and M. Schonlau, “Efficient Global Optimization of Expensive Black-Box Functions,” p. 38.
- [18] H. Dong, B. Song, P. Wang, and Z. Dong, “Surrogate-based optimization with clustering-based space exploration for expensive multimodal problems,” *Structural and Multidisciplinary Optimization*, vol. 57, pp. 1553–1577, Apr. 2018.
- [19] H. Wang, B. van Stein, M. Emmerich, and T. Bck, “Time complexity reduction in efficient global optimization using cluster kriging,” in *Proceedings of the Genetic and Evolutionary Computation Conference on - GECCO ’17*, (Berlin, Germany), pp. 889–896, ACM Press, 2017.
- [20] Mokus, “optimization - Maximizing slow multi-parameter function.”