ABSTRACT

Sound design for products is investigated in this paper as a task that (1) involves different aspects of aesthetics and (2) can use judgments of subjects to measure their quality and guide the design process. Adaptive Bottle shows an exemplified sound design process, which optimizes the input parameters of a physically based sound model using human evaluation in an interactive way. It demonstrates how a 2D parameter space can be searched iteratively. The interaction with the bottle simulates the action of pouring liquid and the emerging sound by this action.

Three optimization experiments have been performed with human subjects, each starting with different initial parameter settings. For each experiment, we observe that the resulting parameter values of the subjects converge to the same region in the parameter space.

1. INTRODUCTION

Previous research on product-sound quality evaluation [1] shows that sound quality can be described as the adequacy of a sound attached to the product. This measure is the combinatory perception different qualities. As in the Bauhaus’ notion, that basic design elements are always linked to their dynamics and only perceived together (line is a cause of a moving point [3]), we follow this principle in product design looking at sound, object and interaction, which are examined by their combined aesthetics. We investigate this by using prototypical design elements: The sounds generated are cartoonified, the object is an abstract bottle like vessel and the interaction is exemplified by a tilting gesture. There is a threefold interrelationship between these aspects: The bottle shape of the object induces a gesture which causes sound that gives feedback and influences the gesture. The changing sound in turn causes a different perception of the object itself (bottle empty/full) that affects the gesture (figure 1).

Sound design as the synthetic generation of sound and aesthetic decision making by controlling the synthesis parameters can be combined with the product and its usage by preference learning in a parameter mapping task [4]. In this respect, physically based sound design offers a novel alternative to recording sounds [5]. We measure adequacy of the generated sounds through judgments of subjects interacting with the object, and optimize the quality of the sounds using statistical methods iteratively.

2. ADAPTIVE BOTTLE OPTIMIZATION

The Adaptive Bottle has a built-in accelerometer, and communicates with the bubbles sound model via the wireless interface of the chip. The acoustic model in [2], based on the use of large quantities of bubbles to represent complex liquid sounds, has been implemented in the Max/MSP environment. It has seven input parameters, which control the statistics of bubbles size, intensity of emission, and rate of bubbles formation. These parameters are used for determining the characteristics of the sound. In this paper, the bubbles size and formation rate parameters are selected for the optimization. The other parameters have been kept constant.

The accelerometer sends 3D orientation information to the computer. By using this information, the tilting angle is calculated. Based on the tilting angle, the volume of liquid remaining in the bottle is calculated. Remaining liquid is used in turn to determine the current bubbles size and the current formation rate. By using the acceleration information, calculated formation rate and bubbles size, and the other synthesis parameters, which are constant,
the physically based sound model generates the bubbles sounds. Figure 2 shows this information flow during the whole pouring action. At the beginning of the interaction, it is assumed that the bottle is full. As the subject tilts the bottle, liquid is poured out, and the bottle becomes slowly empty, depending on the tilting angle. The amount of liquid in the bottle and the emerging bubbles sounds are updated depending on the tilting angle. Intuitively speaking, the size of the bubbles emerging decreases, as the bottle gets empty during the action of pouring. The sound of larger bubbles in pitch is lower than of smaller bubbles. Therefore the bubbles size decreases, as the bottle gets empty.

2.1. Least Squares Optimization

The basic idea behind this optimization is to find the direction and the amount of the optimization step to be made in that direction depending on the evaluation of four sample points, which is supposed to improve the quality of the produced sound. We minimize the unknown preference function by gradient descent. Thereto we model the near surround of the sample point linearly. The direction of a learning step is the gradient of the linear least squares solution. The mathematical formulation of this method according to the Adaptive Bottle problem is as follows:

1. The four data points \( x_i^t \) around the central data point \( x_t \) obtained in the last iteration, which are to be evaluated by the subject are generated. Each of these four points are shifted to the origin, and each dimension is normalized separately.

\[
x_i^t = x_t + r_t \cdot \begin{bmatrix} \cos \left( \frac{\pi i}{2} \right) \\ \sin \left( \frac{\pi i}{2} \right) \end{bmatrix}, \quad i = 1, \ldots, 4
\]

\[
A = \begin{bmatrix} x_1^T \\ x_2^T \\ x_3^T \\ x_4^T \end{bmatrix},
\]

\[
Aw = b,
\]

where \( w \) is the weight vector, \( b \) is the evaluation vector, and \( r \) is the radius. Using the evaluation values \( b \), the optimal \( w \) values are looked for.

2. Squared error is calculated.

\[
E = ||Aw - b||^2
\]

3. The squared error calculated in Equation 4 is minimized, which yields the direction of the gradient in \( w \).

\[
0 = \frac{d}{dw} ||Aw - b||^2.
\]

Taking the derivative, and solving the equation for \( w \) yields:

\[
w = (A^T A)^{-1} A^T b.
\]

4. The new central data point \( x_{t+1} \) is calculated. The learning step \( \lambda \) and the radius \( r \) are updated.

\[
x_{t+1} = x_t + \lambda w,
\]

\[
\lambda_{t+1} = \lambda_t \cdot 0.9,
\]

\[
r_{t+1} = r_t \cdot 0.9.
\]

The decrease of the learning step makes the move in the learnt direction shorter after each iteration. Decreasing the radius causes that the four sample points come closer to each other in the next iteration. Both of these decisions increase the accuracy of the learning. After a certain number of learning steps, the sample points are so close to each other that the difference between the emerging sounds for these samples is not audible anymore. The subject stops the experiment at this point.

2.2. Adaptive Bottle Experiment

The preference learning experiments have been performed on the subjects to test the applicability of such statistical methods for these kinds of optimization tasks. Each subject performed the same experiment three times with three different, preselected initial parameter settings. At the beginning of the experiment, four sample points are presented to the subject around the initial point. One sample point is on the left hand side, one on the right; the other two points are one up and one down (See Figure 3). The subject evaluates all of them one by one in a random
order. The evaluation is supposed to be made in a comparative manner, since after the evaluation, the direction of the learning step is going to be chosen depending on the evaluation, i.e. the combined direction with the highest evaluation rates is calculated. The judgments have values within the interval \([0, 1]\).

The evaluation process is as follows: 4 parameter settings are available for judgment (see equation 2). The subject chooses and listens to them sequentially in an arbitrary order with possible repetition, while performing the action. The subject is able to set the judgment values of all 4 settings at any time, until he is satisfied with the preference ranking. After confirmation, the system will advance to the next 4 setting examples, which have to be judged again the same way. The flow of the experiment is shown in Figure 4.

The user can repeat this evaluation process arbitrarily many times until he / she is satisfied with the quality of the sound. In a typical session five or six learning steps are sufficient. When the subject decides to stop, the trajectory of the whole learning process is shown on a 2D parameter space diagram. The sound corresponding to initial parameter values and the final sound are played as well to show the improvement.

2.3. Experimental Results

The preselected initial points given to the subjects were chosen to be as 1. small formation rate, small bubbles size, 2. large formation rate, small bubbles size, 3. small formation rate, large bubbles size.

The experiments were performed by 15 subjects in total. The summarized results shown in Figure 5 depict only the first and last points of each experiment, in order to show the tendency of each subject. This plot shows all of the three experiments performed by each subject. The bottom right lines depict the results of the experiment with large bubbles size and low formation rate. The bottom left lines are the results of the experiment starting with small bubbles size and small formation rate. Finally the top left lines show the results of the experiment with large bubbles size and small formation rate. The mean of the results for each experiments is shown in Figure 6. In this figure, the mean values of the end points of each subject are depicted for each experiment separately. The starting points are the same as in Figure 5.

As it can also be seen on Figure 5 and 6, the two plots with the small formation rate tend to move in the direction to increase the formation rate. One can also see that the change in the bubbles size increasing for the small bubbles size case, and decreasing for the large bubbles size case, however the main action happens in the vertical direction. The tendency of the curves to move in the vertical direction shows that the formation rate plays a more important role in these experiments than the bubbles size. As a consequence of that, three experiments performed by each subject do not define a closed 2D region in the parameter domain, but rather converge to a certain formation rate region, where the formation rate of the bubbles sound more
realistic. For the cases, where the formation rate value is small, all subjects made moves in the direction of increasing the formation rate, whereas the bubbles size parameter was changed only a small amount compared to the change in the formation rate. However, for the case, where the formation rate is already large, the learning curves generally do not have a common direction.

3. CONCLUSION

We investigated the potential of parameter optimization of a physically based model in product sound design. Based on the notion that the product quality can only be measured when sound, shape and gesture are examined together, we implemented an experimental setup. A local gradient based method on subjective judgments shows common effects over the subjects: The subjective quality is increased step by step and a principal direction in parameter space could be identified. Although the used optimization using a simple update rule, the results encourage to advance to a comprehensive psycho-acoustic evaluation of the matter.

4. FUTURE DIRECTIONS

Statistical methods provide more structure to parameter search problems. However in a 2D domain, random search can converge faster than such an algorithm. Besides, in a 2D domain, in order to make one learning step, four data samples are used. In a higher dimensional domain, this amount increases exponentially, when using two points for every dimension, which makes the problem intractable.

Therefore the model will be improved so that less number of data samples will be needed to make one learning step. Two data samples from the previous evaluation step can be used, and only two new data samples can be presented to the subject. However even this would not solve the problem in a high-dimensional case.

In order to solve the high-dimensional problem, the direction of the learning step should be estimated without evaluating all the data samples around a certain data point. The evaluations in the previous iterations can be taken into account, while presenting to the subject new data samples. Bayesian inference can be built into the model to use the prior knowledge obtained in the previous iterations to calculate a posterior probability of the direction of the next learning step. Hence, a more sophisticated, probabilistic machine learning model will be incorporated to extend the optimization to higher dimensional scale.

5. ACKNOWLEDGMENTS

Thanks to the European CLOSED project (FP6-NEST-PATH measuring the impossible project no. 29085) for funding.

6. REFERENCES


