Integrating Aesthetic and Usability Factors in the Design of Mobile Phones

Dan Nathan-Roberts¹, Yili Liu²

¹Center for Quality & Productivity Improvement, Department of Industrial & Systems Engineering, University of Wisconsin-Madison; ²Industrial and Operations Engineering, University of Michigan

Modern society has come to rely on mobile phones, devices of great functional value and aesthetic importance. Interactive Genetic Algorithms (IGAs) have previously been used to allow individuals to design mobile phones based on aesthetic preference. Here, a multi-objective ergonomic “teammate”, in the form of a Genetic Algorithm (GA) is added to the IGA. GAs and IGAs mimic evolution by iteratively refining a design by having users select designs they find most aesthetically pleasing (IGAs), or using a physical ergonomics model to select ergonomic designs (GAs), and using the selections to create a new generation of designs in a manner similar to natural selection. The web-based experiment had 32 participants, varying button spacing, screen size, and radius of phones over 10 trials (5 with an IGA alone, 5 with an IGA and a GA). The presence of a GA significantly improved the physical ergonomic rating of the final designs, but did not negatively impact the rated aesthetics. IGAs combined with GAs show promise as a tool for designers.

INTRODUCTION

The field of Aesthetic Ergonomics may, to the casual reader, lack an ostensible connection to the larger field of ergonomics referenced in its own title. The ergonomist or human factors professional will be aware however, of the value of a pleasing experience to the quality of life of the user (Norman, 2002), as well as its benefit on their function or perceived function (Tractinsky, Katz, & Ikar, 2000).

Aesthetic engineering is the field of applying quantitative methods to aesthetics. Previously, aesthetic engineering has looked at variety of domains such as mobile phones (Nathan-Roberts & Liu, 2010), bottle design (Kelly, Maheut, Petiot, & Papalambros, 2011), and web page layouts (Bauerly & Liu, 2009). In the past, researchers have focused separately on aesthetic engineering or physical ergonomics (usability). In doing so, designers and researchers removed the very systems nature from our field of systems engineering. Considering usability while considering aesthetics will help realize the goal of true hedonomic design, creating designs that are more desirable and better for the user.

Physical ergonomics in handheld devices is important. For example, mobile phones contribute to cubital tunnel syndrome (also known as “cell-phone elbow”), the second most pervasive peripheral nerve entrapment syndrome (Cutts, 2007).

Other areas of mobile device ergonomics have been investigated as well, including grip span (Kong, Lee, Lowe, & Song, 2007), methods of determining thumb motion and finger force (Ong, 2009), finger abduction speed (Jonsson, Johnson, & Hagberg, 2007), and the effect of screen size on visibility (Hasegawa, Omori, Matsumuma, & Miyao, 2006). Even with this research, mobile devices have significant ergonomics problems beyond cubital tunnel syndrome. “Blackberry thumb,” the overuse of small handheld devices, can cause tendinitis in the thumb (Gordon, 2008). Poor keypad layout and smaller keyboards can slow down data entry (Balakrishnan, Yeow, & Ngo, 2005), cause pain for users with larger hands (Balakrishnan & P. Yeow, 2008), and are often sized inappropriately (Croasmun, 2004).

Research modeling physical ergonomics of handheld devices has been extremely limited. A kinematic model was created to recommend keypad dimensions for a new mobile phone design (Hirotaka, 2003). The authors have previously proposed very simplified physical ergonomics models that attempt to reduce the likelihood of cubital tunnel syndrome, and increase screen legibility (Nathan-Roberts, Beeker, & Liu, 2009).

A multi-objective physical ergonomics model for handheld mobile devices has not been proposed. Following the lead of another multi-objective physical ergonomics model (Brintrup, Ramsden, Takagi, & Tiwari, 2008), a generic model could provide guidance to designers across domains based on the common physical ergonomic problems of handheld devices. The perfect accuracy of such a model is not as important as having a model to provide guidance to designers. The components that are most important to handheld device ergonomics for a multi-objective model would be device length, device width, screen area, and corner radius. A Genetic Algorithm (GA) would be a suitable algorithm to use to make such a model.
Genetic Algorithms

A genetic Algorithm is a computer program which uses an equation to iteratively measure the success of variables in a design space as it explores that design space. Similar to GAs, Interactive Genetic Algorithms (IGAs) iteratively explore a design space, but use a human rater to select the best designs used in the next generation. GAs and IGAs then iteratively evaluate designs against a fitness test, and select the best designs from a group to be used to make the next generation (subject to some mutation and combination of traits similar to evolution). GAs have been used to optimize vehicle suspension components (Alkhatib, Nakaiejazar, & Golnaraghi, 2004), in electromagnetics (Weile & Michielssen, 1997), and other fields.

The inputs to a GA are the independent variables being changed in the design space. The dependent variable in a GA is the fitness test score, which determines if a design is selected for use in the next round. The selections are based on the score of the combination of variables in the fitness test. The number of designs selected depends on the configuration of the GA.

Genetic Algorithms have been combined with Interactive Genetic Algorithms (Brinrup, Ramsden, Takagi, & Tiwari, 2008), but have not been used to combine physical ergonomics and aesthetic engineering. Using an IGA in combination with a GA would allow a balance between a user’s aesthetic score and a quantitative genetic algorithm’s ergonomic score. Combinations can be in parallel or asynchronous. In parallel combination, the GA and IGA run on the same set of designs, their selections are combined, and then iterated using the same method for determining the next generation. While interesting, an asynchronous IGA GA combination may not converge as easily, be difficult to analyze, and is not considered here.

Goals

The goal of this study is to test the ability of an aesthetics focused IGA and physical ergonomics GA to work together to optimize a design based on multiple constraints. The study also seeks to enrich aesthetic ergonomics by combining the sometimes separate fields of aesthetic engineering and physical ergonomics.

METHOD

Participants

All of the participants, except for one were engineering students at the University of Michigan. Of the 32 students, 17 were male, and 15 were female. The mean age was 20 years, with a standard deviation of 1.2 years. Participants used a wide variety of mobile phones; the most common was the Apple iPhone 4 (12 participants). None had been a participant in prior IGA studies. The inclusion criteria were ownership or extensive use of a mobile phone, access to a computer with a high-speed internet connection and 19-inch to 21-inch monitor, and a lack of disabilities that would prevent them from safely using a computer for the duration of the study.

Variables

The independent variables are similar to those used in the exploratory study explained in Nathan-Roberts & Liu (2010), with the addition of the within-subject variable of algorithm-type; an Interactive Genetic Algorithm, or an Interactive Genetic Algorithm combined with a Genetic Algorithm.

The within-subject independent variables that were used to generate a computer rendered mobile phone previously were used here again. The variables were the horizontal and vertical screen dimensions, the horizontal and vertical button spacing, and the roundness (the radius of the outside of the mobile phone and radius of the screen). The radius variable varied the exterior radius and the screen radius simultaneously. The screen radius was one half of the exterior radius value (Figure 1).

![Figure 1. Independent dimension variables in mobile phone design](https://example.com/figure1.png)

Within-subject, the design alternated between being designed with user selection as the only input to the IGA, or with the addition of a genetic algorithm fitness score of the usability of the designs included as well
through a rudimentary ergonomics model. At each trial participants were instructed to select the design(s) that were most aesthetically pleasing.

The primary dependent variables of interest are the selected designs and ergonomic rating of the selections in the last trial where the user could only select one design. Subjective participant preference was also measured through questionnaires after each trial.

The most relevant control variables within the study were the minimum spacing between the screen and outer edge, the minimum spacing between the buttons and the outer edge, and the number of times participants experienced each prompt in the study (five).

**Algorithm configuration**

The algorithms in this study evolved designs over ten iterations. At each iteration, sixteen mobile phones were shown. In the first nine iterations, participants selected their four favorite designs, in the last iteration, they selected their single favorite. Figure 2 shows the web-based interface.

The Genetic Algorithm (GA) was configured to create an ergonomic score for each design. The GA score used all of the independent design variables in the experiment to create a multi-objective equation that simultaneously optimized: phone length, screen area, phone radius, and phone width for usability. The GA rated each objective on a score from one to ten, and the sum of the four scores was considered the ergonomic score of the design. Phone length, screen area, and radius were considered “larger-the-better” design goals based on previous research; the IGA awarded ten points to the maximum possible dimension, and decreased linearly to zero points at the minimum of the range. The phone width design goal was considered “middle-the-best” by the GA equation, giving ten points to the mean of the range of possible widths, and reducing linearly to zero at either end of the range.

When enabled, the algorithm would take the GA score into account in addition to the user selections for each trial. The algorithm would include the top four GA scores as well as the four user selections to determine the parents for the next generation. Each of the top four GA scores were given ten percent of a roulette wheel algorithm to determine the parents of each design for the next generation. Each of the four user selections were also given ten percent of the roulette wheel as well. The maximum percentage of the roulette wheel algorithm a single design could occupy was 20% (if it was one of the four most ergonomic designs and was selected by the user). The remaining designs split the remaining 20% of the roulette wheel algorithm, and the remaining designs split the remaining 20%.

![Image](image_url)

Figure 2. Screen capture of IGA interface with four mobile phones selected, denoted by a darker background.

The IGA and GA were coded in PHP, a popular scripting language, and read into Adobe Flash to render the designs as a webpage. PHP recorded the displayed values and user feedback and saved it to MYSQL databases, an online database system. Variable ranges were set in PHP. The user interface webpage, Figure 2, displayed the participant’s number, the current trial, and a link to proceed to the next trial or survey. A Flash webpage instructed users with a prompt of their goal before each trial.

**Procedure**

Participants were directed to the experiment website via a recruitment email to university departments. Interested participants were qualified via a screening form on the experimental website. Qualified participants were given instructions and a link via email to the study website where they filled out a consent form.

After consent, participants were given an overview of the study and instructed on how to access the study online and use the software tool.

Participants completed ten trials of the IGA software; two practice (one IGA with a GA and one IGA only), and then eight test trials (alternating between the two types of algorithms). After each trial was a short questionnaire, and after all ten trials was a longer debriefing questionnaire. Participants completed the study outside of the lab on a computer meeting the inclusion requirements.
RESULTS

Participant trials were split by the presence of a Genetic Algorithm; half of the trials had only the IGA (n=79), while the other half had an IGA and GA functioning at the same time (n=79). Of the trials rated (n=152), 89.6% said that their “final selection among the best presented in this trial.” Additionally, participants rated 82% (n=124) of the trials “started looking the same” On the trials where it did start looking the same, the mean generation was 5.8 of the 10 generations.

Comparing the first and last generations through paired t-tests, in Table 1, shows that they are statistically significantly different. Table 1 shows that the designs created using an IGA alone are statistically significantly different from designs with a Genetic Algorithm for most variables.

Table 1. Paired t-test p-values comparing the first and last generations, and the difference between IGAs and IGAs with an ergonomic GA.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Buttons</th>
<th>Vertical</th>
<th>Screen</th>
<th>Radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>First to last</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>generation</td>
<td>5.061e-13</td>
<td>5.51e-7</td>
<td>2.2e-16</td>
<td>2.2e-16</td>
</tr>
<tr>
<td>IGAs vs.</td>
<td>0.08531</td>
<td>1.29e-7</td>
<td>3.848e-5</td>
<td>0.1521</td>
</tr>
<tr>
<td>IGA +GA</td>
<td>1.061e-7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*α = 0.01 after Bonferroni correction. *α = 0.00833 after Bonferroni correction.

The presence of a Genetic Algorithm did not statistically significantly alter the aesthetic score of the designs (p-value 0.1367), but did change the ergonomic score, p-value 6.904e-10 (Table 2). The difference in mean values can be seen in Table 3. Subjective aesthetic rating was measured for each final selection on a 0 to 100 scale, with 0 being extremely aesthetically displeasing, 100 being extremely aesthetically pleasing (Table 2).

Table 2. Mean (and SD) aesthetic and ergonomic scores.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Aesthetic score</th>
<th>Ergonomic score</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGAs Alone</td>
<td>80.4 (18.7)</td>
<td>20.85 (7.5)</td>
</tr>
<tr>
<td>IGA with GA as well</td>
<td>75.6 (20.7)</td>
<td>27.2 (4.26)</td>
</tr>
</tbody>
</table>

Table 3. Mean (and SD) values of selected designs in the last generation.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Button spacing</th>
<th>Screen</th>
<th>Radius</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Horizontal</td>
<td>Vertical</td>
<td></td>
</tr>
<tr>
<td>IGAs Alone</td>
<td>6.62 (1.36)</td>
<td>3.22 (1.20)</td>
<td>38.30</td>
</tr>
<tr>
<td>IGA with GA as well</td>
<td>5.30 (1.10)</td>
<td>4.09 (0.74)</td>
<td>38.30</td>
</tr>
</tbody>
</table>

* Bold indicates pairs that are statistically significantly different (α = 0.0083)

DISCUSSION

When combining an IGA with a GA it would be easy to change the parameters unknowingly in such a way that the combined algorithm has a drastically different outcome, is deterministic, or does not converge. The results show that these problems did not manifest themselves. Comparing the first and final generations (Table 1) serves as a good proxy for testing whether or not the algorithm is deterministic when the mean of each independent variable is not the final state as the first generation is random, and therefore centralizes to the mean. Along with testing if the algorithm is deterministic, it is important to test whether or not the algorithm converges. It was clear from subjective questionnaires that the algorithms converged leading to the designs rated as “looking the same.” The data also show that the solution that the algorithm converged to was highly desirable, and that the converged algorithm exceeded the user discriminability in the majority of trials. Overall, the combined IGA-GA algorithm was worked extremely smoothly.

Along with focusing on the mechanics of the algorithm, it is important to see if it achieved the goals of the study, to determine the difference in preference when combined with an ergonomic rater. With most of the independent variables different when the GA was present, it is clear that the GA worked with the user to create more ergonomic designs. The multi-objective equation used in the GA meant that some of the independent variables were “pushed” towards the maximum of their range by “larger-the-better” equations, and some were not. The two variables that were not significantly different were the two that were associated with the “middle-the-best” design goal of the GA, horizontal button spacing and screen width. Had all of the independent variables been significantly different between the algorithms, or had they all remained the same, it would be harder to tell if the combined algorithm was appropriately weighting the potentially competing values of its two inputs. Similarly, the lack of significantly lower aesthetics ratings when the GA was combined with the IGA indicates that the GA did not overpower the user in sharing responsibility for the final design. Overall, the combined algorithm pushed designs to be taller, and have more rounded edges.

The physical ergonomics GA used here is an incredibly simplified model of physical ergonomics, which was not validated through real-world or 3D experimentation. The model is used here illustratively to test the ability of the combined algorithm to enhance the safety of the final design. Even though the algorithm is rudimentary, the resultant designs are more physically
ergonomic. Similarly, a major limitation of this study was that participants could not hold the phones.

Building a combined IGA and GA algorithm for ergonomics has not been done before. It is therefore important to test the concept in domains like mobile phones where ergonomics and aesthetics are well studied.

The study demonstrates that IGAs combined with GAs can be used for design. A key difference between this study and other work is the strict focus on aesthetics and ergonomics instead of less concrete terms such as “liking” (Brintrup, Ramsden, Takagi, & Tiwari, 2008). Greater detail about IGA-GA configurations can be found in Nathan-Roberts, in-press, but was omitted here due to space constraints. Using IGA-GA combinations has the potential to have a large impact in the device design market. Future work on the types of constraints that a GA can provide with an IGA will be important.

Future work should refine the multi-objective fitness function used by Genetic Algorithm, test this methodology in other domains, such as blood glucose meters, where users are not as experienced, and explore combining this work with rapid prototyping technology. Additional physical ergonomics research is needed to enhance the realism of the designed algorithm. The multi-objective fitness function can also be used to include design constraints for the designs, such as required design envelope size.

REFERENCES


