Eye on the Prize: Using Overt Visual Attention to Drive Fitness for Interactive Evolutionary Computation

Tim Holmes
Royal Holloway, University of London
Department of Psychology
Egham, Surrey TW20 0EX
+44 (0) 1784 443 704
t.holmes@rhul.ac.uk

Johannes Zanker
Royal Holloway, University of London
Department of Psychology
Egham, Surrey TW20 0EX
+44 (0) 1784 443 521
j.zanker@rhul.ac.uk

ABSTRACT
Interactive Evolutionary Computation (IEC) has been applied to art and design problems where the fitness of an individual is at least partially subjective. Applications usually present a population from which the preferred individuals are selected before the usual evolutionary operations are performed to produce the next generation. Large population sizes and numbers of generations impose significant demands on the user. This paper proposes that selecting by means of eye movements could reduce user fatigue without sacrificing quality of fitness assessment. In the first experiment, an eye-tracker is used to capture fixations and confirm the reliability of such a measure of attention as a fitness driver for subjective evaluation such as aesthetic preference. In a second experiment, the robustness and efficiency of this technique is investigated for varying population sizes, presentation durations and levels of fitness sampling. The results and their consequences for future IEC applications are discussed.

Categories and Subject Descriptors
J.4 Social and Behavioural Sciences – *Psychology*

J.7 Computers in Other Systems – Consumer products

General Terms
Algorithms, Design, Experimentation, Human Factors, Performance, Reliability.

Keywords
Attention, Design/synthesis, Eye-movements, Fitness Evaluation, Genetic Algorithm, Interactive Evolutionary Computation, Visual Perception.

1. INTRODUCTION
The ability of evolutionary algorithms to search complex, multi-dimensional solution spaces and locate highly fit individuals has been applied to art and design with varying degrees of success. When the fitness is primarily objective in nature, such as some combination of weight, size and cost of components constrained by properties such as stability and surface area, the algorithms have successfully generated feasible and innovative designs [1, 2]. When the fitness includes a large subjective component, such as in evolutionary art [3, 4] and consumer based design [5], the most successful approach has been to use humans for the fitness evaluation, giving rise to the field Interactive Evolutionary Computation (IEC) [6].

IEC thus allows for true phenotypic fitness assessment, where the overall fitness cannot be reduced to some mathematical combination of the fitness of its parts. However, the introduction of a human immediately limits the performance of such algorithms by slowing down the processing of a single generation, limiting the number of generations and restricting the number of individuals which can be presented for evaluation by the human observer [6, 7]. Attempts have been made to reduce user fatigue through fitness interpolation or including some degree of machine learning capability which enables the evolution to proceed without the involvement of the human at every generation [8]. Such methods can reduce the rate of fitness improvement in the population, with sporadic jumps in fitness occurring when information from the user is obtained, but with relatively small gains when the fitness is estimated [7]. This suggests that more value could be obtained by increasing the speed and quality of the feedback from the human observer which would in turn facilitate more frequent sampling from the user, over and above the development of improved fitness estimation processes.

The human visual system provides a well designed mechanism for the processing of complex object related information [9]. The subjective assessment of individuals in a population can be compared with 'psychophysical' paradigms, whereby stimulus properties are systematically adjusted according to the subjective responses of a human participant based on percepts [10]. We demonstrated that evolutionary algorithms can be applied to research into aesthetic perception where participants are requested to select their preferred stimulus - which is then used to generate a new population of offspring. This method generates results comparable with those of more established psychophysics methods [11].

Visual search paradigms using eye-trackers, where eye-movements and their associated latencies are used to research the parallel processing of visual features and contextual effects [9], provide a means of rapidly accessing quantitative measures of subjective preference which can then be integrated into real-time interactive evolutionary algorithms. Wolfe's guided visual search model proposes that in order to facilitate locating an object in space, humans create spatial maps for each of the visual features they are interested in (e.g. brightness, orientation, size, colour, etc.) and use conjunctions between those features in space to shift the spotlight of their attention [12]. The Rizzolatti, et al. pre-motor theory of attention suggests that the shift of the attentional spotlight is result of preparing to make an eye-movement (saccade) to an object of interest [13]. The common ground shared by these models is that eye-movements are a reliable indicator of shifts in overt attention. Search accuracy is improved by instructing participants to make rapid saccades to the objects they are attending to [14], and the fixation points on the screen can be used as indicator of the location of overt attention. The experiments described below establish the use of fixation positions and the duration of overt attention directed towards an individual as a robust measure of perceptual fitness in guided visual search experiments.

2. EXPERIMENT 1 - AESTHETIC PERCEPTION

Experiment 1 builds on a wealth of previous research into whether there exists an aesthetic preference for the Golden Ratio (1:1.618) [15, 16]. In a previous experiment we have shown that an evolutionary algorithm can be used to generate populations of rectangles consisting of mutations of the rectangle deemed "most beautiful" by the participant in the previous generation [11]. In this experiment, the selection of the preferred rectangle based on a key press is replaced by selection based on the amount of time spent fixating on the individual rectangles which is captured using an eye-tracker. The results are compared with data from a previous experiment using the manual selection method.

2.1 Representation

The total surface area of the rectangle was held constant to maintain constant size on the retina, a necessary control in such experiments [16], and so a simple genetic representation of just one side of the rectangle was sufficient. The length was encoded as a 9 bit Gray coded binary integer [17], which corresponded to the displayed horizontal length of the rectangle. The smallest displayable length is one pixel, and as such the physical lengths were continuous to this degree of accuracy.

2.2 The Algorithm

A simple (1+λ) evolutionary strategy was used [18] i.e. the population of size $n = 8$ comprised the preferred individual from the previous generation (parent) plus $\lambda = n-1$ offspring generated using mutation only. In this case mutation was via a simple bit-flip operator at a random location, which occurred with a probability of 0.1. All trials used a randomly generated starting population.

The algorithm and all interfacing software were written in MATLAB.

2.3 Experimental Set-up

2.3.1 Stimuli

Populations of 8 rectangles were rendered and displayed using a Cambridge Research Systems ViSaGe (Visual Stimulus Generator). Stimuli were displayed on a 48cm CRT Monitor at a distance of 57cm from the participant. Constant size ($630mm^2$) and orientation (portrait or landscape) was maintained for all rectangles in the population. The maximum range of ratios displayed was from 1:1 to 1:4. This was chosen to ensure that the Golden Ratio was not the mid-point of the range. Rectangles were displayed in a radial fashion to counter positional bias on the screen, with a small random positional jitter to ensure that participants could not exploit alignment of images in their assessment – see Figure 1. The position of the parent rectangle on the screen was randomly selected for each generation. Each participant ran through 6 trials comprising 10 generations, for both landscape and portrait orientation. Participants were instructed to search for the most aesthetically pleasing rectangle.

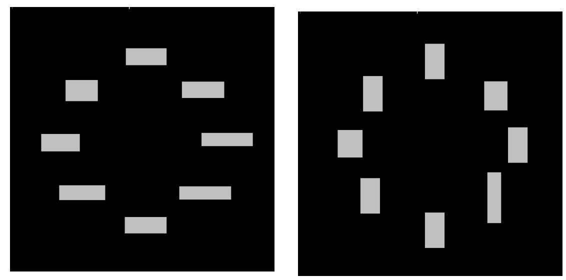

Figure 1. Sample stimuli, showing starting populations of landscape and portrait rectangles.

2.3.2 Timing

Participants were initially presented with a black screen with a central fixation cross for 1000ms. The population of rectangles was then presented together for a duration of 750, 2500 or 5000ms, after which a black screen was presented for 3000ms before the fixation cross was displayed. This ensured that participants returned their gaze to the centre of screen between each generation, and allowed time for any after-images caused by the high contrast stimuli to fade. In the manual selection condition, there was no limit on the presentation duration and participants were simply directed to respond when they identified a preference.

2.3.3 Eye-tracking

A Cambridge Research Systems 50Hz Video Eye-Tracker was used, which enables gaze location to be sampled every 20ms. Fixations were defined as periods of 100ms or more, within a zone which enclosed the rectangle. Due to the size of the stimuli we used a 25mm tolerance window which eradicated noise caused by positioning eye-movements, small saccadic eye-movements within a fixated region and eye-movements which pass over a rectangle but do not stop to fixate on the rectangle.

2.3.4 Fitness Assessment

After presentation of the stimuli, the gaze positions on the screen were analyzed at 20ms intervals. Any positional information which did not fulfill the criteria for a fixation (see above) was removed as well as any fixation which lay outside of the fixed zone containing each rectangle. The total amount of time spent fixating in each zone was then calculated and the rectangle displayed in the zone with longest total fixation duration was used as the parent for the next generation. In the event that the participant did not fixate on any of the 8 rectangles, a parent for the next generation was selected at random.

2.4 Results

Not all participants were able to locate and fixate on a preferred rectangle for the 750ms presentation, so this data has been excluded. Figure 2 shows the average ratio and standard deviation for the population at generation 10 for both landscape and portrait rectangles at the 2500ms and 5000ms presentation durations, which all four participants were able to complete. As can be seen, there is no significant difference between the evolved ratios based on manual selection (key pad) or fixation duration. The results were stable for both the 2500ms and 5000ms presentations. As expected, the variation in the results can be controlled through the presentation duration, with the variance within participants at 5000ms presentations being similar to that in the earlier manual selection experiment [11], where the mean rectangle selection time was 5583ms (st.dev = 2314ms).

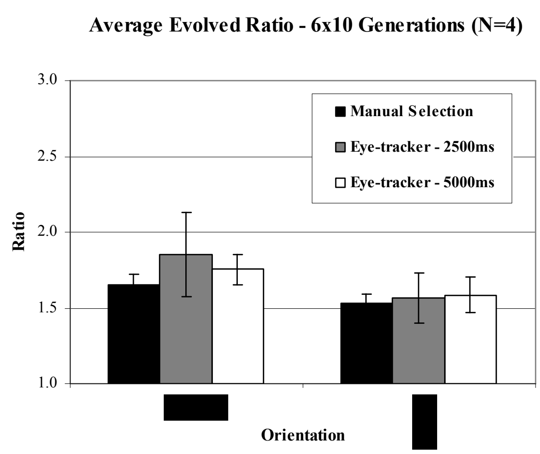

Figure 2. Average evolved ratio for manual and fixation duration based selection (2500ms and 5000ms) of most aesthetically pleasing rectangle. Results are shown, with their standard deviations, for both landscape and portrait rectangles. N = 4.

Figure 3 shows average evolved ratios for the full range of exposure durations for the one participant who was able to complete the experiment at the shorter durations of 1000ms and 750ms. Again, the variance appears to increase with reduced exposure duration.

2.5 Discussion

The results show that, for simple phenotypes such as these, fitness selection based on fixation duration can be used to evolve individuals which are comparable with those evolved by manual subjective interactive selection, with significantly shorter presentation durations. The stability of the results decreases with reduced presentation durations, with some untrained participants being unable to maintain stable fixations in the experiment at very short durations (750ms). This suggests that some degree of training is needed if shorter durations are to be used. The ability of a practiced participant to complete the experiment at durations of 1000ms or less suggests that eye-tracking might present us with the possibility to increase the number of individuals evaluated and generations used in IEC, which is important for larger solution spaces. To validate this hypothesis, a more complex phenotype with a correspondingly larger solution space is needed, and is explored further in experiment 2 below.

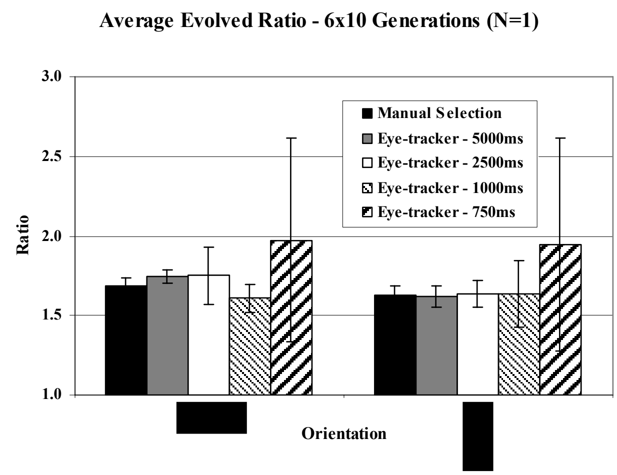

Figure 3. Average evolved ratio for manual and fixation duration based selection of most aesthetically pleasing rectangle. Results are shown, with their standard deviations, for both landscape and portrait rectangles. N = 1.

The increased variability in the results at shorter presentation times might be due to a distractor effect. As the population begins to converge, individuals that deviate significantly from the population average tend to "pop-out". Research has shown that such items tend to attract attention [19] suggesting that it might be preferable to apply a lower, or even zero weighting, to the first 100ms. However, the noise introduced by such involuntary attentional shifts may be of interest, particularly in product design, where the target packaging is intended to be a distractor itself, particularly when surrounded by competing brands. An alternative approach to limiting distractor noise would be the introduction of dynamic display durations, where the exposure time is inversely proportional to the similarity of the targets. Similarly, the introduction of dynamic fitness weights proportional to the difference between an individual and a converging population could be used to promote distractor effects.

3. EXPERIMENT 2 - MULTI-DIMENSIONAL PERCEPTION

In order to establish that eye-tracking can be used for aesthetic evaluation of more complex phenotypes it is important to understand the effects of increasing the solution search space and varying the exposure duration. Experiment 2 explores these effects with an enhanced genetic algorithm which allows larger population sizes to be sampled for presentation as well as introducing crossover and partial replacement to avoid premature convergence. To control for noise resulting from differences in participants' aesthetic preference, participants were given a specific target to search for and were instructed to look for the closest individual in the displayed stimuli.

3.1 Representation

The phenotype comprised a 3x2 array of four possible shapes – square, circle, triangle or blank. Each of the shapes could be presented as a white shape on a black background, or a black shape on a white background. Figure 4 shows the 8 possible occupants of each location in the grid. Thus the chromosome comprised 6 such genes, with a total length of 18 bits, giving 2^{18} = 262,144 possible phenotypes. Figure 5 shows an example of a randomly generated phenotype and the target phenotype. Participants were shown an example of the target and told to look for the individuals which were most like the target.

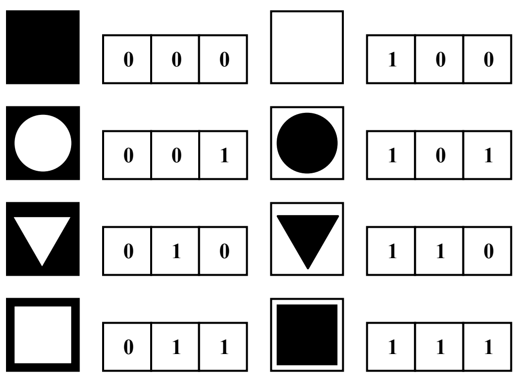

Figure 4. Binary representation of the 8 possible alleles for each of the 6 grid positions (genes).

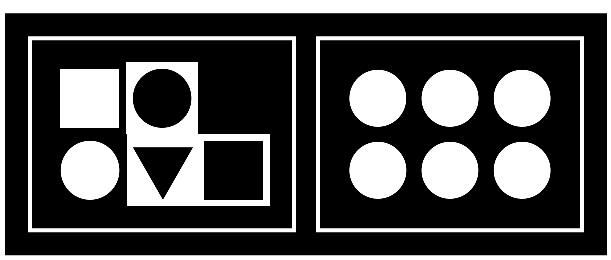

Figure 5. Random population member (left) and target population member (right).

3.2 The Algorithm

For this experiment a simple genetic algorithm [17] employing both crossover and mutation was used to reduce the risk of premature convergence with the increased phenotype complexity. Simple bit-flip mutation was used, with a constant probability of 0.05 and single-point crossover with a constant probability of 0.75. Roulette wheel selection was used with the probability of selection directly proportional to the fitness of the individual. All trials used a randomly generated starting population.

The algorithm was enhanced to facilitate partial replacement, which has been shown to reduce the risk of premature convergence [20]. A further enhancement to the algorithm was the sampling of the set of individuals to be presented from the total population. This has the benefit of allowing for an increased population size while maintaining a relatively small number of presented individuals, and therefore reducing user fatigue, however it requires a means of estimating the fitness of the population members which are not presented.

The algorithm and all interfacing software were written in MATLAB.

3.3 Experimental Set-up

3.3.1 Stimuli

Samples of 8 individuals were rendered and displayed using a Cambridge Research Systems ViSaGe (Visual Stimulus Generator). As before, stimuli were displayed on a 48cm CRT Monitor at a distance of 57cm from the participant. The physical size of an individual on the screen was 300mm^2. Sampling was with replacement, meaning that the same individual could be displayed more than once. This was intended to limit contextual effects since the fitness is always relative to the other individuals presented. Individuals were displayed in a radial fashion as before, but without any positional jitter. Each participant completed 10 trials of 20 generations each, with presentations of 1, 2 and 4 samples per generation.

3.3.2 Timing

As in the previous experiment, participants were initially presented with a black screen with a central fixation cross for 1000ms. Samples of 8 individuals from the population were then presented together for display durations of 750ms and 1500ms after which a black screen was represented for 3000ms before the fixation cross was represented.

3.3.3 Eye-tracking

As in the previous experiment. Fixations for the zone enclosing the entire phenotype were captured; fixations on individual features within the phenotype were not distinguished.

3.3.4 Fitness

When subjective fitness information is only available for a sample of the population it is necessary to estimate the fitness scores for the remaining population members. Previous attempts at this in IEC have included interpolation of the fitness based on Euclidian distance of the individual from the most and least preferred member of the presented sample [7]. Eye-tracking provides the ability to capture fitness values for all members of the presented sample.

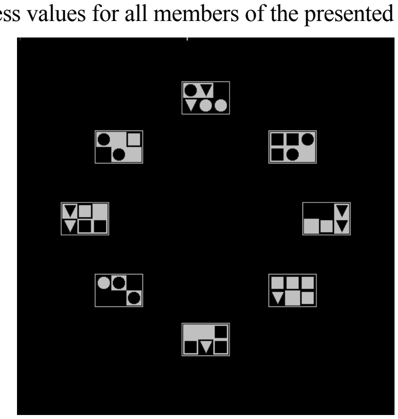

Figure 6. Example of presented display.

Let x_{ij} represent the ith population member in generation j. Let the fitness of the ith individual in the jth generation be denoted by $f(x_{ij})$ and be defined as the average amount of time spent fixating on that individual over all presentations for the jth generation. Then

$$f(x_{ij}) = T(x_{ij}) \,/\, N(x_{ij}) \Big/ \sum_{i} T(x_{ij}) \,/\, N(x_{ij}) \tag{1}$$

Where $T(x_{ij})$ is the total amount of time spent fixating on the ith population member in generation j, and $N(x_{ij})$ is the total number of presentations of the ith population member in generation j (1).

Once all samples had been presented for one generation, the fitness of each un-presented individual was estimated. Let $H(x_{ij},x_{kj})$ be the hamming distance the ith (un-presented) individual and the kth (presented) individual. This is calculated for x_{ij} for all un-presented x_{kj}, and the method of least squares is then used to estimate the equation of the line which best describes the points $(H(x_{ij},x_{kj}),\ f(x_{kj}))$. This equation can then be used to estimate the fitness of the un-presented individual as follows:

$$\hat{f}(x_{lj}) = \alpha H(x_{lj}, x_{lj}) + \beta \tag{2}$$

Since the hamming distance of an individual with itself is always zero, β gives the estimated fitness of the un-presented lth population member (2).

This process is repeated for all un-presented population members before the selection and reproduction stages of the algorithm.

3.4 Results

The performance of the algorithm can be evaluated in terms of its ability to evolve an individual which matches the target, but a more relevant measure for IEC is to look at the population as a whole. This is especially important when the population is sampled for presentation to the participant, since a target might be evolved without the participant ever being aware of it. Results shown are for an individual trained participant.

3.4.1 Population Level Results

The proximity of each individual to the target can be assessed genotypically in terms of hamming distance, or phenotypically in terms of visual similarity. This latter measure is appropriate to the success measurement of a subjective interactive evolutionary algorithm. The target comprises an array of 6 white circles on a black background (target characteristic), and so the visual similarity can be assessed in terms of the number of white circles on a black background in each individual. Figure 7 shows a typical example of the target characteristic increasing in prevalence over time. In all runs the average prevalence of the target characteristic can be seen to increase, suggesting that the fixation duration based fitness is successful in evolving the desired pattern.

Population size (n), presentation duration and the number of presentations (samples) per generation each affected the performance of the algorithm, as can be seen in Figure 8. For n=8, increased sampling through multiple presentations of the same 8 individuals improves the performance of the algorithm. 1500ms presentations provided a small improvement in the average similarity to the target, although there is little difference between the two presentation durations. In all cases variance reduced with increased presentation duration. For n=96, increased sampling produced mixed results, showing a 50% reduction in the variance but no improvement in the prevalence of the target characteristic. Increased presentation time produced more significant improvements in performance, with corresponding reductions in variance as well. It is worth noting that over 20 generations there seems to be no real gain from the increased population size, however, the prevalence does not show signs of saturation which can be seen in the smaller population size suggesting that further improvements in the average fitness could be achieved by allowing the experiment to run for more generations.

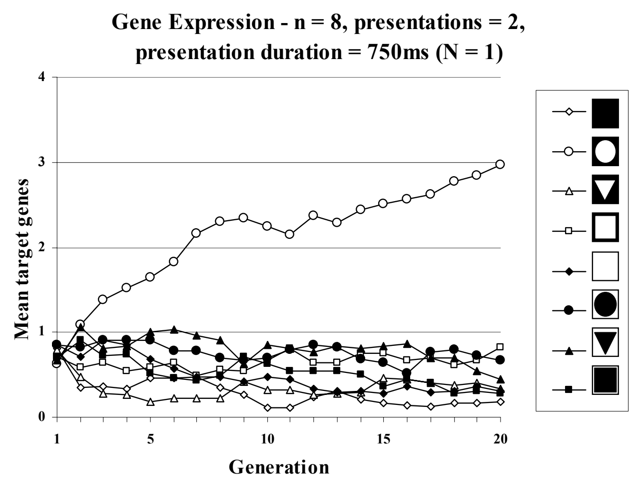

Figure 7. Typical expression of phenotypic characteristic over 20 generations. Averaged over 10 trials, N = 1.

3.4.2 Fittest Individual Results

As already mentioned, the ability to achieve the target is a somewhat irrelevant measure of the performance of this algorithm, especially in the event that the evolved target is not presented to the participant. However, the proximity of the fittest individual to the target at each generation is of interest, as it is a measure of the algorithm's ability to be directed towards to the target by the participant.

Figure 9 shows that in all cases increasing the presentation time and sampling improves the fitness of the fittest individual in the population. Although increased population size improves the fitness of the fittest individual in the population, which is simply the result of a larger initial sample, it does not affect the rate of fitness improvement.

3.5 Discussion

The above results confirm that, whilst fixation duration is an appropriate fitness measure for use in IEC, its effectiveness is subject to the duration and number of presentations of sample individuals, as well as the underlying population size. The improvements resulting from longer presentation durations are most likely the result of a longer final fixation, since participants tend to stop scanning the screen when they find an individual they consider to be closer to the target than they have seen before. This is comparable with a fixated gaze while pressing the selection button in a more traditional manual selection paradigm and results in a higher relative fitness score for the last fixated individual which, of course, manifests itself through an increase in the prevalence of its genes in the next generation. This increased relative fitness seems to compensate for the increased signal-noise ratio in larger populations where fitness estimation is needed and consequently this effect virtually disappears for the smaller population. This suggests some non-linear weighting of the fixation durations might be appropriate.

The reduction in variance resulting from increased sampling is not surprising, since this also improves the reliability of the information, although the benefits were not fully realized in the larger population, suggesting sampling without replacement might be a better technique to use.

A simple linear extrapolation technique was used to assign fitness to individuals that were not presented. The effectiveness of this was constrained because the fittest individual presented represents a maximum fitness threshold for the generation. As such, fitter individuals which are not presented are currently not properly rewarded which might explain the relatively low population averages; future versions of the algorithm will correct this along with a tournament selection, which is more appropriate in when the fitness function is not known [21] as well as introduction of reproduction operators more suited to evolutionary design

The best fitness results show that in the search for a single optimal solution, larger populations improve performance. This is a well documented result [6, 17], but this experiment clearly shows that it is not necessary to present the entire population in IEC, and with the quality of fitness data received from eye-tracking, it is might not be necessary to invoke complex machine learning algorithms to exploit the information collected from a small sample of presented individuals.

4. GENERAL DISCUSSION

IEC is reliant on the ability of the user to assess the proximity of a presented individual to some notional optimal individual. In the situation where this optimal individual can be visualized and retained in memory, this paper has shown that the use of oculomotor information, as indicator of observer attention, can be at least as effective as manual selection, with the added advantage of minimizing user fatigue through shorter presentation durations and sampling from a large underlying population.

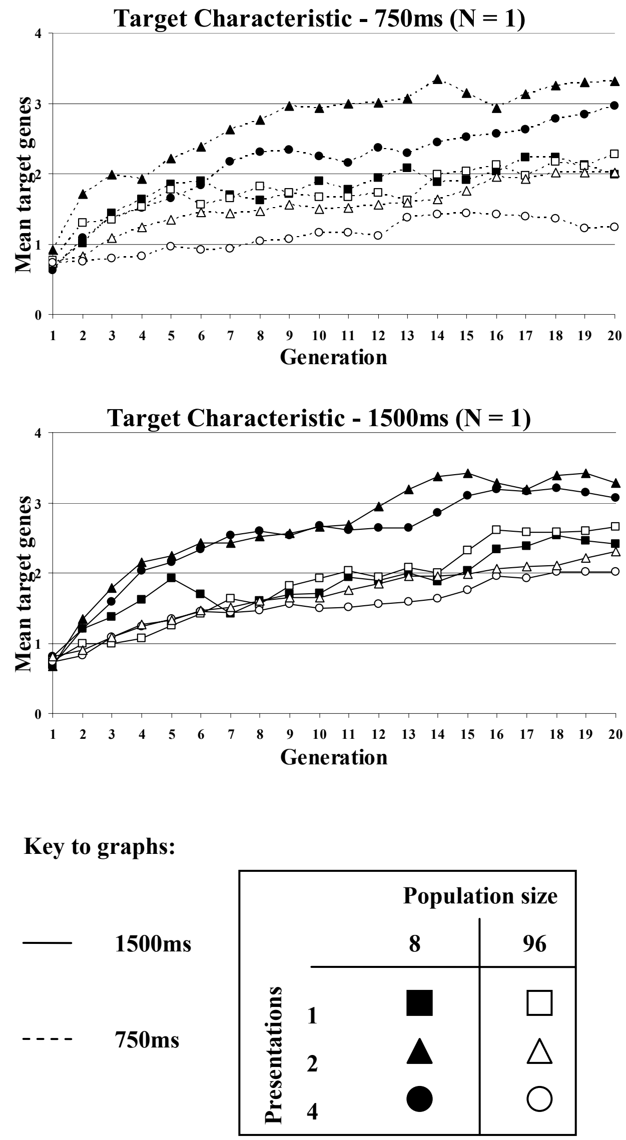

Figure 8. Expression of phenotypic characteristic for population sizes 8 and 96, presentation durations of 750ms and 1500ms and 1, 2 and 4 presentations per generation. Averaged over 10 trials, N = 1.

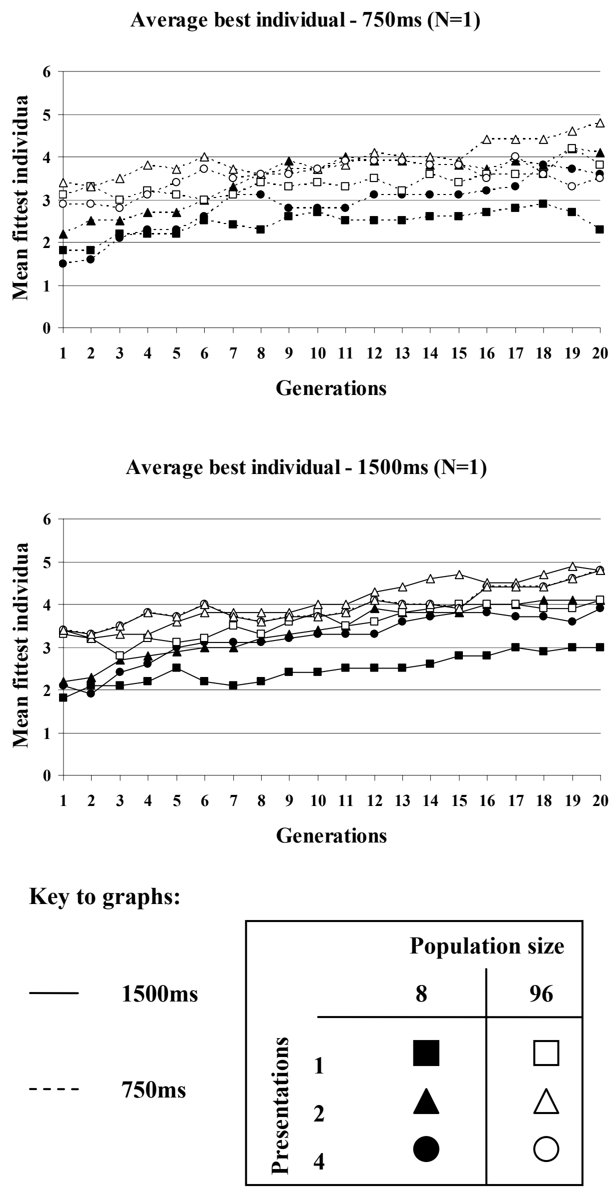

Figure 9. Expression of phenotypic characteristic for population sizes 8 and 96, presentation durations of 750ms and 1500ms and 1, 2 and 4 presentations per generation. Averaged over 10 trials, N = 1.

Eye-trackers, such as the one used in these experiments, provide the opportunity for further development of the subjective fitness function to include other measures such as the order of fixations, the number of times the gaze returns to a specific individual and even changes in the pupil size during observation of the individuals. The sequence order of fixations could be used to eliminate "pop-out" attentional effects as well as to enhance the fitness of the last location to be fixated, as this often represents the preferred choice of the user. The number of returns to an individual could be useful assessing the strength of preference for one individual relative to the presented population as return fixations suggest the individual is being compared with other population members by the user. It has been suggested that pupil diameter might provide an insight in to the strength of the user's preference [22]. This would require all luminance and adaptation effects to be removed before it could be used, and as such requires further research before it could be applied.

Repeated sampling of an individual improves the quality of the information obtained from it, as can be seen from the results of Experiment 2. The current algorithm does not utilize this in estimation of fitness and so assumes the same degree of noise in all fitness values. As sampling increases, so too does the potential to exploit this in the algorithm and implementation of such an enhancement might provide a control for some of the less consistent assessments of fitness which inevitably occur due to user fatigue, distractions or even blinking.

An underlying assumption in the performance measurement of this algorithm is that proximity to target can be measured in terms of the number of white circles against a black background which are present in the evolved phenotype. Whilst this is a valid measure, it became clear during executions of the experiment that participants tended to adopt more general strategies during the early generations to drive the evolution towards the target. These strategies included searching for individuals with the most circles, regardless of colour, and searching for individuals with the most white shapes on a black background. This suggests that IEC such as this might be able to exploit this through the introduction of a more sophisticated measure of similarity, perhaps by analyzing the fitness within each dimension and using this information to apply additional weight to the fitness scores for those phenotypes which show uni-dimensional convergence. This approach would potentially result in a more rapid localization of the search, but probably only makes sense in a uni-modal solution space. In multi-modal solution spaces, some form of Pareto optimization [17, 23] could be implemented.

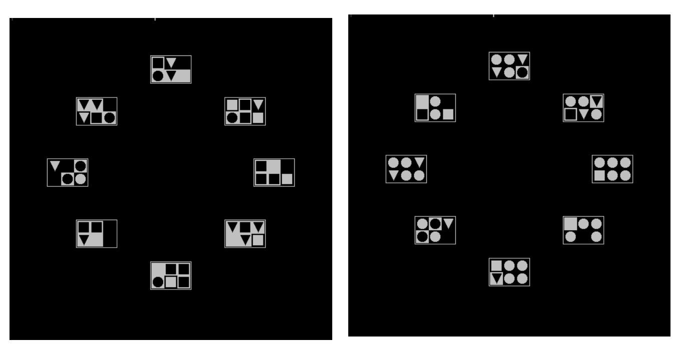

Figure 10. Example of a generation 1 sample presentation (left) and generation 20 sample presentation (right).

Figure 10 shows an example of a generation 1 presentation sample and a generation 20 presentation sample. As has already been mentioned, the task of quickly searching the display for a definite preference becomes increasingly difficult as the population begins to converge. This has the effect of increasing the noise in the user's fitness assessment and explains why the rapid gains made in early generations flatten out in later generations. Additionally, in the example shown, the target individual already existed in the underlying population, but was never presented to the user. These two problems could be addressed by introducing dynamic sampling and dynamic presentation durations. As the population variance reduces, the sampling rate needs to increase. As the sample variance reduces, the presentation duration needs to increase. The effects of these enhancements will be the focus of future research.

4.1 Conclusion

Experiment 1 successfully employed an eye-tracker in the evolution of an aesthetically pleasing rectangle through IEC. Experiment 2 used an "evolve to target" paradigm to explore the performance of an evolutionary algorithm which uses fixation data captured by means of an eye-tracker to estimate the fitness of more complex individuals and assess the effects of varying the presentation duration, sample size and underlying population. Together these results firmly establish the appropriateness of the integration of attentional fitness drivers to the field of Interactive Evolutionary Computation and highlight some of considerations which must be addressed as part of any real-world applications of this technique.

5. ACKNOWLEDGMENTS

We would like to thank Szonya Durant, Andrew Meso, Robin Walker, Chris Watkins and the anonymous reviewers for their helpful comments and useful suggestions. This work is supported by the Engineering and Physical Sciences Research Council, UK and Procter & Gamble. Grant number 05002329.

6. REFERENCES

[1] Bentley, P. J. 1996. Generic Evolutionary Design of Solid Objects using a Genetic Algorithm. Huddersfield, University of Huddersfield. PhD.

[2] Funes, P., and Pollack, J. 1999. Computer Evolution of Buildable Objects. Evolutionary Design by Computers. ed. P.J. Bentley. pp.387-404. San Francisco: Morgan Kaufmann.

[3] Todd, S., and Latham, W. Evolutionary Art and Computers. Academic Press, Inc. 1994.

[4] Sims, K. 1991. Artificial Evolution for Computer Graphics. Computer Graphics, 25, 319-328.

[5] Machwe A. T., and Parmee I. C. 2006. Integrating Aesthetic Criteria with Evolutionary Processes in Complex, Free-form Design - an Initial Investigation. IEEE Congress on Computational Intelligence, Vancouver, Canada.

[6] Takagi, H. (2001). Interactive evolutionary computation: Fusion of the capabilities of EC optimization and human evaluation. Proceedings of the IEEE, 89 (9), 1275-1296.

[7] Quiroz, J. C., Louis, S.J., and Dascalu, S.M. 2007. Interactive Evolution of XUL User Interfaces. Proceedings of Genetic and Evolutionary Computation Conference 2007, University College London.

[8] Machwe,A. T, Parmee I. C,. 2006. Introducing Machine-learning within an Interactive Evolutionary Design. Environment. Design 2006. Dubrovnik, Croatia.

[9] Findlay, J.M and Gilchrist, I.D. 2003. Active vision: The psychology of looking and seeing. New York: Oxford University Press

[10] Snodgrass, J.G. 1975. Psychophysics. In: Experimental Sensory Psychology. B Scharf. (Ed.) pp. 17-67.

[11] Holmes, T., and Zanker, J. 2007. Evolving the Golden Ratio: A new method for an old question of aesthetic perception, Perception. Pion 36: 107

[12] Wolfe, J.M. 1994. Guided search 2.0 – a revised model of visual search. Psychonomic Bulletin and Review, 1, 202-238.

[13] Rizzolatti, G., Riggio, L., Dascola, I., and Umiltà, C. 1987. Reorienting attention across the horizontal and vertical medians – evidence in favor of a premotor theory of attention. Neuropsychologia, 25, 31-40.

[14] Klein, R.M. and Farrell, M. 1989. Search performance without eye-movements. Perception and Psychophysics, 46, 476-482.

[15] Livio, M. 2002. The golden ratio: the story of phi, the extraordinary number of nature, art and beauty. London: Headline Review.

[16] Green, C.D. 1995. All That Gltters: A Review of Psychological Research on the Aesthetics of the Golden Section, Perception. Pion, 24, 937-968.

[17] Goldberg, D. E. 1989. Genetic Algorithms in Search, Optimization and Machine Learning. Boston, Addison Wesley Longman Inc.

[18] Beyer, H.-G, and Schwefel, H.-P. 2002. Natural Computing 1: 3–52

[19] Joseph, J.S., and Optican, L.M. 1996. Involuntary attentional shifts due to orientation differences. Perception and Psychophysics,58(5), 651-665

[20] Goldberg D.E., and Deb K. 1991. A comparative analysis of selection schemes used in genetic algorithms. Proceedings of Foundations of Genetic Algorithms FOGA-I. ed. G. Rawlings. pp.69–93. Morgan Kaufmann.

[21] Miller, B.L., and Goldberg, D.E. 1995. Genetic Algorithms, Tournament Selection, and the Effects of Noise, IlliGAL Report No.95006.

[22] Pallez, D., Collard, P., Baccino, T., and Dumercy, L. 2007. Eye-Tracking Evolutionary Algorithm to Minimize User Fatigue in IEC Applied to Interactive One-Max Problem. Proceedings of Genetic and Evolutionary Computation Conference 2007, University College London.

[23] Deb, K. 2001. Multi-objective Optimization Using Evolutionary Algorithms. Chichester: Wiley.