Disguise: A Game that Evaluates Visualization Algorithms

Nafees U Ahmed
Stony Brook University
Stony Brook, NY
nuahmed@cs.stonybrook.edu

CHI 2014, Apr 26 - May 01 2014, Toronto, ON, Canada
ACM 978-1-4503-2474-8/14/04.
http://dx.doi.org/10.1145/2559206.2580100

Abstract
'Disguise' is an arcade action survival game that was designed with an intent to explore the possibilities of using purpose driven games as an alternative method to existing techniques of visualization evaluation. We designed Disguise as a proof of concept that such technique can indeed provide significant help for both scientific and information visualization research.

Author Keywords
Games for a Purpose; Evaluation; Visualization; Human Based Computation;

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction
Visualization, in simple words, is a process that transforms abstract data into a visual form. A person looking at the picture can get a quick understanding of the underlying data. If a visualization can help a person communicate well with the data, we can consider it a success. Measuring such success directly is still considered hard or impossible. Because, we are yet to fully understand how the human brain and its visual sensors system works. As such, effective evaluation of

Color Blending and Transparency Order Perception

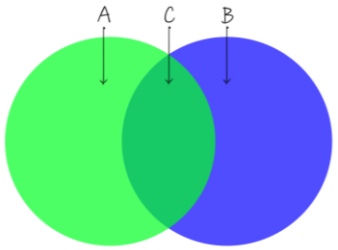

Figure 1 Composite color C is produced from the blending between A (green) and B (blue). The human eye perceives A to be on top of B.

Color Blending: Color blending is the process of creating a composite color from two semitransparent colors. From Figure 1, layer A (Green) is placed on top of layer B (Blue). Color blending algorithm produces composite color C.

Transparency Order Perception: Given the composite color, can a person perceive the true ordering or the original layers? From Figure 1, looking at color C, can a person guess if A was actually on top or B?

visualizations is considered one of major challenges or obstacles for the research community [7] and we still consider user evaluation to be the most trust worthy method. Arranging enough participants for such a user study to get statistically significant number of data points is hard, costly and sometimes infeasible. Crowd sourcing user study tasks can help us significantly raise the number [5]. But, the motivation being money only, it produces its own limitations [6]. In an effort to search for a suitable alternative, we explored the possibility of using Human Computation (HC) [9] methods. HC relies on the idea that, in many tasks, human brain is far superior to the computers that we have now. We can achieve much more complex goals by bringing in humans into the computational loop. One very effective method of HC is transforming the task into a purpose driven game. Scientific research have already seen several glowing success in use of purpose driven games [3][4]. At the time of our design, we were yet to see a similar effort for visualization field. We decided to explore how a game based evaluation would work and how successful they can be compared to previous methods. As a test case, we decided to create a game that would evaluate performance of different algorithms for blending between two semitransparent layers of color. The result was the game titled 'Disguise'. We built disguise as a proof of concept that games are also viable method for visualization evaluation.

Purpose of the Game

The goal of the game 'Disguise' is to evaluate the performance of the selected color blending algorithms with respect to their measure of transparency order perception.

Figure 2 Screenshot of the game.

Gameplay

The game tells a story of a planet that is under attack from *Intruders*. Intruders are small disc looking saucers that move around the space fast and eventually explode with an intent to destroy the planet. The job of the player is to defend against them. To help the cause, two things are provided. First, a weapon that can be fired to destroy the intruders. Second, *Collectors* that help gather valuable information about the attack. *Collectors* are big circular devices that stay still and monitor the activities.

Core Mechanics

Intruders move around the screen before exploding and damaging the collectors. The player can destroy them by firing on them. But, the player only collects points when he can successfully destroy an intruder disc on top of any collector. Intruders, being sneaky, takes disguise (becomes semitransparent) and sometimes

Actions and Responses

Fire an intruder

- If the intruder is on top of a collector, destroy intruder and earn points.
- If the intruder is below a collector, damage collector.
- If the intruder is away from collector, destroy intruder but earn no points.

Fire a collector

- Damages collector.

Fire in empty space

- Nothing happens

Do nothing

- No positive reward. The intruders damages the collector once they time out and explode.

choose to hover below the collectors. Trying to fire an intruder when they are below, damages the collector. The player has to master the skill of making right guess about the ordering by looking at their color and act fast.

Design Process

Defining the objective

In visualization research, the user study task to evaluate blending algorithm involves two overlapping circles with composite color showing in the middle. The participant has to choose color he thinks is on top. We had to design a game that encourages the players to do the same tasks with entertainment as the only motivation.

Design of game characters

Making a game that does the given task in larger scale poses a very strong restriction on the visuals. To ensure correctness of the experiment, we are only allowed to use single colored semitransparent circles without any form of visual deformation. Any design decision we make has to make sure that it doesn't violate this given constraint. We created collectors and intruders with simple colors and circular shape. Intruders are distinguished by their small size and fast random movement. Color of the intruders and collectors are randomized to cover the experiment space.

Design of rules that achieve the objective

The primary action the player can do is to fire his weapon (left click on the screen). We collect data every time he clicks. We designed the game rules to ensure his actions conform to our requirements. We do so by assigning positive or negative rewards to each possible actions. Actions and Responses section included on the left of this page explains how our design requires the player to guess if the intruder circle is on top of the collector circles. The player can only excel by learning this skill. Otherwise all collectors gets destroyed and the game ends.

Designing mechanics to keep player engaged

The core mechanics ensure correctness of data collection. Other mechanics ensure entertainment and thus engagement from the players. We experimented with following options,

- To keep the players in the flow, we introduced controlled increase in difficulty through levels. We controlled the level of challenge using these parameters: transparency level, speed of intruders, rate of generation of intruders, explosion rate, attack variety and color.
- Consecutive right guesses on the same collector creates and increases score multiplier.
- Leaderboards for with social plugin.
- Shuffle collectors for convenient arrangements.
- Soothing audio tracks.

Results and Evaluation

We launched the game and invited players through social media. After several weeks of keeping it live we analyzed the collected data. We can discuss the success of the game design from two perspective: as a game and as a visualization evaluator.

As a game

First 15 days of game play engaged 261 unique players, generating close to 30,000 data points. The number can be considered very insignificant compared to what an engaging game is able to achieve. We collected feedbacks from the players about what they felt about our game, what seemed missing and how they can be improved. We found players generally intrigued by the idea that they are helping a research and found the game mechanics to be unique. But, they also criticized the predictability of gameplay and lack of strong re-playability value. This was expected, because, our current design still lacks balancing through rigorous play testing. Also, the possibility of introducing more engaging game mechanics still remains unexplored.

As a visualization evaluation

The game was very simple in design and we had a reasonably short term of data collection. But, even then data collected through our game showed some strong signs. We were able to reproduce faithfully previously proven results. Also, due to high quality of data from motivated gamers, we also deduced some newer facts. This proved our claim that games can be a very useful technique for visualization evaluation. The fact that even a simple prototypical design was able to produce such results should encourage the researchers to explore this path more. For further information refer to our publication related to the game 'Disguise' [1].

Acknowledgements

We thank Prof. Klaus Mueller (Stony Brook University) for advising us throughout the whole research from the very beginning. We thank Ziyi Zheng for helping in writing our article about game driven evaluation. We also thank 'The Piano Guys' [8] for their magical sound tracks that kept the players immersed and keeping this game designer motivated through the hard working hours of development.

References

[1] Ahmed, N., Zheng, Z., and Mueller, K. Human computation in visualization: Using purpose driven games for robust evaluation of visualization algorithms *IEEE Transactions on Visualization and Computer Graphics*, 18, 12 (2012), 2104-2113.

[2] Chen, C. Top 10 Unsolved Information Visualization Problems, *IEEE Comput. Graph. Appl.* (2005), 25, 12-16.

[3] Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popovic, Z., and Foldit players. Predicting protein structures with a multiplayer online game, *Nature* (2010), 466, 756-760.

[4] Eterna. http://eterna.cmu.edu.

[5] Heer, J. and Bostock, M. Crowdsourcing graphical perception: using mechanical turk to assess visualization design, *Proceedings of SIGCHI* (2010), 203-212.

[6] Ipeirotis, P.G., Provost F., and Wang J. Quality management on Amazon Mechanical Turk, presented at the Proceedings of the ACM SIGKDD Workshop on Human Computation (2010), 64-67.

[7] Plaisant, C. The challenge of information visualization evaluation, *Proceedings of the working conference on advanced visual interfaces* (2004), 109-116.

[8] The Piano Guys http://thepianoguys.com/

[9] Von Ahn, L. Human computation, *Design Automation Conference* (2009), 418-419.