

Mining Human Shape Perception with Role Playing Games

Golam Ashraf

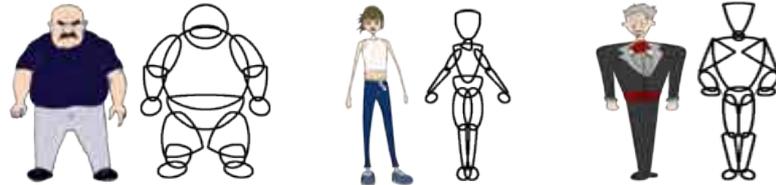
National University of Singapore
Blk AS6, 13 Computing Drive
Singapore 117417
+65 65167348
gashraf@nus.edu.sg

Yong Peng Why

National University of Singapore
Blk AS4, 9 Arts Link
Singapore 117570
+65 65168857
psywyp@nus.edu.sg

Md. Tanvirul Islam

National University of Singapore
Blk AS6, 13 Computing Drive
Singapore 117417
+65 65167348
tanvirulbd@gmail.com



Abstract—‘Games with a purpose’ is a paradigm where games are designed to computationally capture the essence of the underlying collective human conscience or commonsense that plays a major role in decision-making. This human computing method ensures spontaneous participation of players who, as a byproduct of playing, provide useful data that is impossible to generate computationally and extremely difficult to collect through extensive surveys. In this paper we describe a game that allows us to collect data on human perception of character body shapes. The paper describes the experimental setup, related game design constraints, art creation, and data analysis. In our interactive role-playing detective game titled Villain Ville, players are asked to characterize different versions of full-body color portraits of three villain characters. They are later supposed to correctly match their character-trait ratings to a set of characters represented only with outlines of primitive vector shapes. By transferring human intelligence tasks into core game-play mechanics, we have successfully managed to collect motivated data. Preliminary analysis on game data generated by 50 secondary school students shows a convergence to some common perception associations between role, physicality and personality. We hope to harness this game to discover perception for a wide variety of body-shapes to build up an intelligent shape-trait-role model, with application in tutored drawing, procedural character geometry creation and intelligent retrieval.

Keywords—Game with a Purpose, Shape Perception, Human Computation, Web Games

1. INTRODUCTION

Increasing sophistication in programming languages, parallel hardware, social networking and machine learning algorithms have enabled powerful new media applications. However, many problems are still hard to model on computers without the aid of intelligent human input. Common examples are in segmentation and recognition of objects, subjective opinions on similarity of patterns, perception and emotion related feedback.

The human computation approach [13] was developed to harness intelligent human input to train sophisticated AI algorithms. In this method, a participant is usually assigned a small fragment of a large problem. By assigning task fragments to many users it is possible to solve a large and computationally untraceable problem within a reasonable amount of time. Various techniques are being developed for making this participation enjoyable. The most effective technique among these is computer games. ‘Games with a purpose’ (GWAP) [14], is a genre of serious games, developed with the focus of efficient and entertaining transformation of research data collection into game mechanics. By immersing the player in the game puzzle, and tying the players’ success to providing quality inputs, the hope is to extract higher quality data than surveys.

The Entertainment Software Association has reported that more than 200 million hours are spent each day playing computer and video games in the U.S. By the age of 21, the average American spends more than 10,000 hours playing such games [9]. This clearly shows the potential of games with a purpose to channel some of this human effort for doing useful things like solving computational problems or gathering data for training artificial intelligence algorithms.

Depending on the nature of the research question, the game design, namely: core mechanics, and user experience model; varies significantly. In this paper we design a GWAP that allows us to collect data on human perception of character body shapes. We describe the experimental setup, related game design constraints, art creation, and data analysis. The key contributions are in the sharing of the specific transformation of our research question, as well as preliminary analysis of some feedback data on character design, which is also quite relevant to game design.

We have created an interactive role-playing detective game titled Villain Ville. Players are asked to characterize different versions of full-body color portraits of three villain

characters. They are later supposed to correctly match their own character-trait ratings to a set of characters represented only with outlines of primitive vector shapes. Our purpose is to find out how the basic shape decomposition compares to the full visual design in terms of perception by a layman audience. We hope to harness this game to discover perception for a wide variety of body-shapes to build up an intelligent shape-trait-role model, with application in tutored drawing, procedural character geometry creation and intelligent retrieval.

2. RELATED WORK

The use of online collaboration to solve large goals dates way back to the 1960s. More recently, Web 2.0 technologies have eased the creation of online forums, blogs, surveys and collaborative document creation. Wikipedia is an example of large-scale collaborative knowledge creation, moderation and sharing. It is possible to mine information on characters from this semi-structured knowledge source, but the information is mainly restricted to well known characters from published media. Amazon Mechanical Turk employs a cloud computing architecture to allow publication of paid surveys, that they coin as Human Intelligence Tasks (HITs). Each survey can be denominated a desired number of HITs and disseminated to registered human workers, within a designated time window. While this system of micro-payment and some basic filtering of input outliers, seem an attractive data collection option to researchers, the caveat is that it is hard to get motivated data for complex questions, for typically a few cents per completed task. The open mind framework [11,12] is another worldwide endeavor for leveraging on human ability to train computers. Volunteers participate by providing answers to questions computers cannot answer (e.g. “What is in this image?”), aiming to train recognition and segmentation models. Unfortunately, these inputs are also vulnerable to random white noise from unmotivated participants.

“Interactive Machine Learning” [4] provides live feedback to training inputs to indicate how well the regressive machine learning algorithm is performing. This feedback loop is perhaps more effective than the above methods in reducing poor inputs from humans.

GWAP systems are different because they are designed to be enjoyable, ensuring motivated data collection, thus minimizing the probability of random white noise. According to human-computer interaction researchers, it is very important to introduce the elements of enjoyment and fun in user interfaces [10, 20]. Many researchers suggested that introducing game-like interfaces can be a key to keep the user motivated. Games and role-playing activities are becoming popular modes of learning for children [7].

Among the notable GWAP systems today, are The ESP Game [15], a.k.a. the Google Image Labeler. Players provide meaningful, accurate labels for images on the Web as a side effect of playing this game; for example, an image of a man and a dog is labeled “dog,” “man,” and “pet.” Instead of using computer vision techniques that don’t work well in the presence of missing information, poor lighting or new data, ESP harnesses human labels (gathered in playful

manner) for powering its image search. As a result, we often get surprisingly intelligent answers to our image queries in Google. People don’t play these GWAPs because they think they are doing some useful work, rather than do it solely for enjoyment. Other GWAPs include Peekaboom [16], which locates objects within images; Phetch [17], which annotates images with descriptive paragraphs; and Verbosity [18], which collects commonsense facts in order to train reasoning algorithms.

The above mentioned GWAPs are essentially multiplayer-online games, where usually two players are randomly pitted against each other. The research objective is usually to extract labels common to both opponents, but the game mechanic may instead reward uncommon labels to generate excitement. We design our GWAP of extracting human perception on basic body shapes as a single-player role-playing game instead. This demands more careful attention to game mechanics, art and story development to keep the player immersed.

3. RESEARCH QUESTION

Basic shapes like triangles, circles and squares are so well understood, that even a textual/verbal description of structures in terms of these shapes elicits a natural visualization in our brain. Basic shapes play an important role in character design drafts. For example, artists use shape scaffolding to pre-visualize the final form, using basic shapes to represent each component or part [1]. Apart from establishing the volume and mass distribution of the figure, these shapes may also help portray a certain personality, as is widely seen in stylized cartoon drawings. Depending on the art style, primitive shapes may become less apparent with the addition of details; e.g. clothes, accessories, and hair for humanoid figures.

In this paper, we investigate the role of basic shapes in layman perception of character traits. The player rates fully fleshed out characters sampled by three broad body-type categories. The player also needs to rate characters drawn with only primitive shapes. This allows us to capture the comparative perception labels between the two versions. We also want to capture perception labels for as many different primitive body shapes as possible.

4. GAME DESIGN

Villain Ville is a role-playing game, where the player dons the role of a detective and pursues missions to bust a criminal ring. Each mission involves talking to people, gathering clues, challenging the villain to an action puzzle, and finally rating the villain (see Fig. 1).

We realize that rating characters with labels is a tedious task. Our full label set contains 10 attributes per character, each with a High/Medium/Low/Null value. The attributes traits can be grouped under *Role (Action, Charm, Skill)*, *Physicality (Power, Metabolism)* and *Personality (Intelligence, Diligence, Sociable, Diplomacy, Emotional)*. In order to make the game exciting, we turned around the “rating” of primitive shape characters, to a “matching” problem. The player has to select the most appropriate primitive character to match an

attribute-value trait, randomly ordered from his/her own set of ratings for the 3 main fully fleshed out villain characters. This way, we can return a “right”/“wrong” result for every choice to satisfy the players need to know whether he/she got it “correct”. However, whatever choice the player makes, it is valuable to the research question, as it captures what really appears to be the most appropriate primitive character shape (with no other design stimulus) for a given trait label.



Fig. 1. Rating full visual designs of villain characters

The primitive shape matching part kicks in as a finale. The boss character transforms his cronies into these shapes as a final attempt to escape the long arm of the law. The player needs to successfully match his/her 30 rated attributes to scrolling outline drawings (by shooting with the cross-hair cursor in Fig. 2), and get the majority answers “right”.



Fig. 2. Matching of trait labels to basic shape outlines

We had to balance the entertainment value of the game as well as the research constraints. For example, since the player is not rating too many characters, it may not be hard to remember the physical shape of the characters and cheat during the outline-matching finale. Furthermore, we wanted to ensure that the player does not get fatigued with the rating, so we inserted significant story elements and action puzzles to reduce the players’ short term memory of the rated characters. We avoided time-bound ratings for the full visual designs, but incorporated this in the matching finale. This was to ensure that we did not get any missed rating data, but allowed a timer level-up of increasing difficulty in the matching part, to break the monotony of matching 30 labels back to back. We did lose some information for a few labels for some players, but this helped us overcome general player boredom, since nearly all players finished the mini-game.

5. CHARACTER ART

It was important to decide how to sample and expose the characters to the players, because of the large number of combinations of shape and proportion of different body parts. We broadly divided the character body shapes into 3 types: fat, thin and stereotypical (to the portrayed roles), as shown in Fig. 3. The rows represent three villains: a thug, an assassin, and a mastermind. We deliberately kept all the clothing and accessories similar across the three versions of each villain, so that the visual detailing differences did not confound our experimental results.

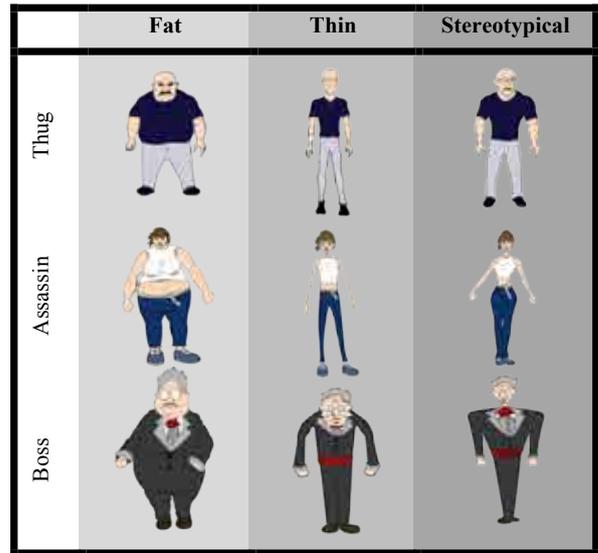


Fig. 3. Body type classification for Main Villains

To ensure generalization of the ratings, participants who were exposed to more than one main characters of the same body-shape had their attribute ratings for these body-shapes averaged.

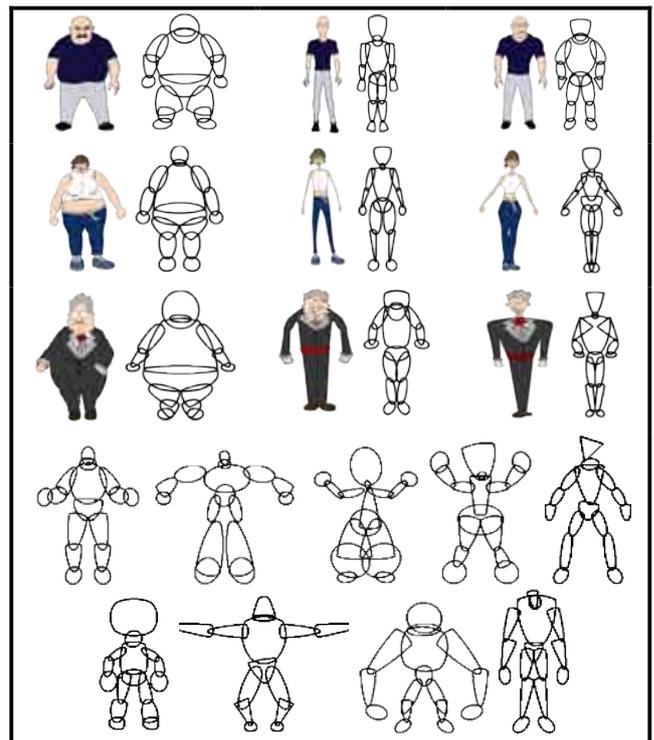


Fig. 4. Primitive Body shape decompositions for Main Villains and secondary characters taken from popular action figures

Fig. 6 illustrates the associated primitive shape decompositions for the main and secondary characters used in the matching mini-game. The secondary characters add challenge to the game, as well as get free ratings for all “wrong” choices. This way, we can systematically introduce unlimited variety and number of main and secondary characters for shape-attribute labeling by players.

6. RESULTS

Supervised game play sessions were conducted at a secondary school and a university lab on two different groups of 50 teenagers (Group A) and 50 undergraduates (Group B). The sessions were supervised as this was a pilot release of the game, and we wanted to ensure that all the data was collected reliably. The facilitators did not guide the players, but simply sat with them or walked around in the lab to ensure that nobody faced any problems. Only after the sessions, did the facilitators take feedback and suggestions for improvement. Most participants welcomed the idea and design, but suggested a more active mechanic for better engagement. This was especially true for teenage boys who said they preferred more action and less role-playing. Nonetheless, we were successful in obtaining conclusive perception feedback and present the analysis of data collected from both groups.

Fig. 5 shows a comparison on matching and rating labels gathered from Group A. The frequency distribution is much higher in the matching graph, as we allowed multiple matching sessions per game, triggered in the event of “losing” in the matching mini-game. Nevertheless, we find similar trends between the two graphs. Specifically, individual traits show almost similar distinctions for both the primitive shape and full visual design versions of the character, as shown in more detail in Fig. 6. The results for the rating and matching of a character trait are converted to percentage (over the total number of people who rated the certain trait) for a better comparison. The percentage obtained is rounded off to nearest whole number for comparison. A difference in 25% and above between matching and rating a certain trait is seen as significant different results while a difference in 10% and below is seen as significant similar results.

Let us first present a slightly verbose description of the results shown in Fig. 5 and 6, with the example of the fat assassin character. In Fig. 6, the rating and matching results for the Fat character variation is similar in terms of ‘Charm – High’ (7% VS 5%), ‘Intelligence – Low’ (67% VS 57%), ‘Beauty – High’ (7% VS 8%), ‘Sociable – High’ (7% VS 4%), ‘Diligence – Medium’ (20% VS 15%). This shows that both the full design and primitive shape versions exhibit several similar traits. However, the results significantly differ in terms of ‘Beauty – Low’ by 0.53 on normalized scale (13% VS 66%). In this case, it shows that a fat, primitive body shape tends to be more strongly associated to ‘ugly’ compared to a full-design character. The results also differs in terms of ‘Built – High’ by 0.28 on normalized scale (87% VS 59%) and ‘Action – Low’ by 0.26 on normalized scale (87% VS 61%). In both cases, the full-design character was more strongly associated to the respective traits compared to its corresponding body shape.

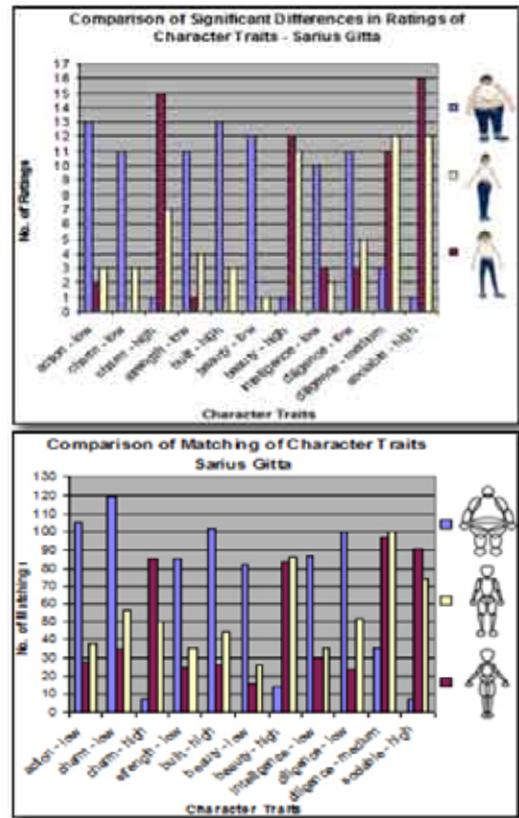


Fig. 5. Comparison of significant differences in ratings and matching of key character traits for the variations of the assassin character

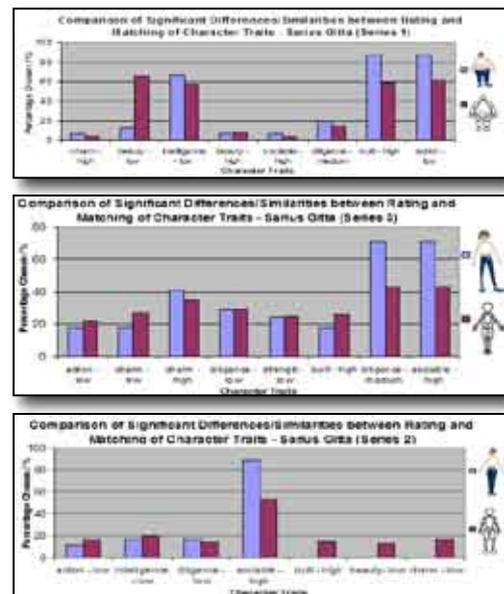


Fig. 6. Comparison details of significant differences in perception of full design and primitive shaped assassin character

We now present more generic analysis from the secondary school (Group A) perception data. Fig. 7 shows a combined perception label measure for the three character types: fat, thin and stereotypical, for all the main character primitive outline shapes (see top 3 rows of Fig. 4). We assigned the following numbers for the attribute values (Null:0, Low:1, Medium:2, High:3). Visually we can immediately see differences in action, beauty and build between Fat body types and the other two types in Fig. 7. Also, the ratings for strength, intelligence, and emotional attributes are quite

similar for all the three body types. We submitted this data to SPSS statistical analyses using GLM Repeated Measures. The following attributes display significant statistical difference with respect to the Stereotypical character group: Fat body-type (Action, Charm, Build, Beauty, Diligence, Sociable) and Thin body-type (Action).

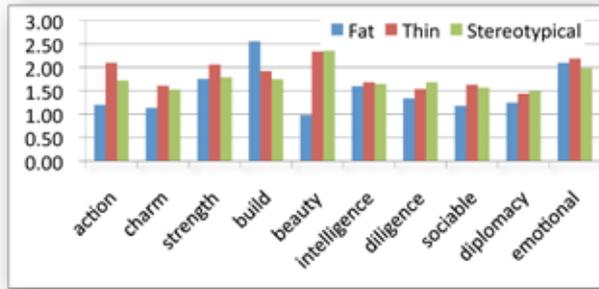


Fig. 7: Perception values for Fat, Thin and Stereotypical Primitive Shaped Characters

We present the same analysis as above in Table 1, except that this was computed from undergraduate university students (Group B). The standard deviation in the labels is low, showing that we can safely rely on the mean values. With respect to the Stereotypic group, we find that the following attributes have statistically significant difference: Fat body-type (Action, Charm, Skill, Metabolism). All these attributes got significantly lower ratings compared to the stereotypic body type. Note that we decided to supplant a more intuitive term (Build vs. Metabolism) for the secondary school students. Both terms imply the same thing, except that the label values are opposite, giving us a congruent analysis; i.e. we expect lower metabolism for higher build value.

Table 1. Mean (Standard Deviation) Ratings by Attribute and Bodyshape Condition (numbers in bold are significantly different within each row)

	Fat	Thin	Stereotypic
Action	1.14 (0.20)	2.13 (0.19)	2.25 (0.21)
Charm	1.36 (0.18)	1.88 (0.17)	2.05 (0.19)
Skill	1.50 (0.19)	1.96 (0.19)	2.05 (0.20)
Strength	2.00 (0.22)	2.00 (0.21)	1.90 (0.23)
Metabolism	0.96 (0.19)	2.08 (0.18)	1.80 (0.19)
Intelligence	1.32 (0.18)	1.92 (0.17)	1.50 (0.19)
Diligence	1.05 (0.21)	1.58 (0.20)	1.10 (0.22)
Sociable	1.27 (0.20)	1.71 (0.19)	1.60 (0.21)
Diplomacy	1.18 (0.18)	1.38 (0.17)	1.55 (0.19)
Emotional	1.36 (0.21)	1.42 (0.20)	1.35 (0.22)

7. CONCLUSION

We have implemented a character perception capturing game, disguised as a detective role playing game. The similarities in the rating and matching results for the main characters show that both the full-design and primitive body

shapes exhibit the same traits. Slight differences in results may be due to the player being able to judge better through the full-design character by their facial expression and texture details.

We have presented only preliminary analysis of a few body shapes in this paper. We are currently in the process of sampling from a wide array of existing character art in comic books, film, game and casual web art. We hope to expose a more representative sampling of character shapes through our game, and thus collect more refined information on how basic shapes affect the final perception.

We are also refining the design of Villain Ville, to make the rating and matching activities more closely knit to the economy of the game. So a player needs to do these activities regularly to earn more money or get smarter, so that they are able to clear certain barrier points in the game. This will help players perform the exercise more diligently, and generate more research data. We are also working on better reward and ranking to motivate the players, as well as provide a more immersive, richer world with well-designed characters. We are confident that this game is going to be a great instrument for collecting perception data for character design, that will help our ongoing work in intelligent drafting, automatic labeling and shape queries.

ACKNOWLEDGEMENT

We would like to thank Vania Stephanie, Ng Li Yin, Sally Tan, from Nanyang Girls High School, Singapore and Rebecca Chen from NUS School of Computing for implementing the game and conducting the surveys. This research is funded by the NUS Cross Faculty Fund, WBS: R-581-000-084-646.

REFERENCES

- [1] Camara, S., All about techniques in drawing for animation production, 1st ed. 2006, Barron's Education Series, Inc. ISBN: 0764159194
- [2] Chao, D. Computer games as interfaces. *Interactions* 11, 5 (Sept.–Oct. 2004), 71–72.
- [3] Chao, D. Doom as an interface for process management. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Mar. 31–Apr. 5). ACM Press, New York, 2001, 152–157.
- [4] Fails, J.A. and Olsen, D.R. Interactive machine learning. In *Proceedings of the Eighth International Conference on Intelligent User Interfaces* (Miami, Jan. 12–15). ACM Press, New York, 2003, 39–45.
- [5] Giles, Jim. Internet encyclopaedias go head to head. *Nature*, 438:900–901, 2005.
- [6] Locke, E.A. and Latham, G.P. *A Theory of Goal Setting and Task Performance*. Prentice Hall, Englewood Cliffs, NJ, 1990.
- [7] Malone, T.M. What makes things fun to learn? Heuristics for designing instructional computer games. In *Proceedings of the Third ACM SIGSMALL Symposium and the First SIGPC Symposium on Small Systems* (Palo Alto, CA, Sept. 18–19). ACM Press, New York, 1980, 162–169.
- [8] Malone, T.M. Heuristics for designing enjoyable user interfaces: Lessons from computer games. In *Proceedings of the Conference on Human Factors in Computing Systems* Gaithersburg, MD, Mar. 15–17). ACM Press, New York, 1982, 63–68.
- [9] Richards, C. Teach the world to twitch: An interview with Marc Prensky, CEO and founder Games2train.com. *Futurelab* (Dec. 2003); http://www.futurelab.org.uk/resources/publications_reports_articles/web_articles/Web_Article578.

- [10] Shneiderman, B. Designing for fun: How can we design user interfaces to be more fun? *Interactions* 11, 5 (Sept.–Oct. 2004), 48–50.
- [11] Stork, D.G. and Lam C.P. Open mind animals: Ensuring the quality of data openly contributed over the World Wide Web. In *Learning from Imbalanced Data Sets: Papers from the AAAI Workshop (Technical Report WS-00-05)*. (Austin, TX, July 30–Aug. 1). American Association for Artificial Intelligence, Menlo Park, CA, 2000, 4–9.
- [12] Stork, D.G. The Open Mind Initiative. *IEEE Intelligent Systems & Their Applications* 14, 3 (May–June 1999), 19–20.
- [13] vonAhn, L. Human computation, *Proceedings of the 4th international conference on Knowledge capture*, p.5-6, October 28-31, 2007, Whistler, BC, Canada
- [14] vonAhn, L., Laura Dabbish, *Designing games with a purpose, Communications of the ACM*, v.51 n.8, August 2008
- [15] vonAhn, L. and Dabbish, L. Labeling images with a computer game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Vienna, Austria, Apr. 24–2)*. ACM Press, New York, 2004, 319–326.
- [16] vonAhn, L., Liu, R., and Blum, M. Peekaboom: A Game for locating objects in images. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Montreal, Apr. 22–27)*. ACM Press, New York, 2006, 55–64.
- [17] vonAhn, L., Ginosar, S., Kedia, M., and Blum, M. Improving image search with Phetch. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (Honolulu, Apr. 15–20)*. IEEE Press, New York, 2007, IV–1209–IV–1212.
- [18] vonAhn, L., Kedia, M., and Blum, M. Verbosity: A game for collecting common-sense knowledge. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Montreal, Apr. 22–27)*. ACM Press, 2007, 75–78.
- [19] Waterman, Donald A. *A guide to expert systems*, Reading, Mass. [u.a.] Addison-Wesley 1986, The teknowledge series in knowledge engineering, ISBN:0-201-08313-2
- [20] Webster, J. Making computer tasks at work more playful: Implications for systems analysts and designers. In *Proceedings of the SIGCPR Conference on Management of Information Systems Personnel (College Park, MD, Apr. 7–8)*. ACM Press, New York, 1988, 78–87.



Golam Ashraf received his PhD in Computer Engineering from NTU, Singapore. He teaches game development, animation and interactive media at School of Computing, National University of Singapore. His research interests are in real time graphics, computational aesthetics, multimedia analysis, and pedagogical game design. Website: <http://www.comp.nus.edu.sg/~ashraf/>



Yong Peng Why received his PhD in Psychology from St. Andrews, UK. He teaches at Faculty of Arts and Social Sciences, National University of Singapore. His research interests include relationship between personality and cardiovascular regulation during psychological stress and recovery, and relationship between subjective appraisal, objective stress conditions and cardiovascular stress arousal. Website: <http://profile.nus.edu.sg/fass/psywyp>



Md. Tanvirul Islam received his Bachelors degree in Computer Science and Engineering from Bangladesh University of Engineering and Technology in 2008. He is currently a Research Assistant at School of Computing, National University of Singapore. His research interest is in machine learning, data mining, multimedia retrieval and graph theory.