Proceedings of the ASME 2016 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2016
August 21-24, 2016, Charlotte, United States

DETC2016-60081

IMPRESSIONIST: A 3D PEEKABOO GAME FOR CROWDSOURCING SHAPE SALIENCY

Houpu (Hope) Yao
Aerospace Engineering
Arizona State University
Tempe, Arizona, 85287
Email: hope-yao@asu.edu

Max Yi Ren*
Mechanical Engineering
Arizona State University
Tempe, Arizona, 85287
Email: yiren@asu.edu

ABSTRACT

Designers often express their intents (e.g., on product functionalities and semantics) through shape features. Therefore, collecting such "salient" features from existing shapes and learning their associations with design intents will enable efficient design of new shapes. However, the acquisition of saliency knowledge from a large shape collection has not been accomplished. This paper investigates a gamification approach to this end. In addition, we propose to validate a derived saliency map by its corresponding shape recognition accuracy through crowdsourcing. This allows a comparison across existing and the proposed saliency acquistion and computation methods. Currentl results show that the proposed method achieves statistically similar recognition accuracy to existing saliency data on a standard shape database, indicating that various saliency maps are equally valid according to the proposed saliency definition. Nonetheless, the saliency data obtained through the proposed game consistently produces reasonable viewpoints across shapes, outperforming existing curvature-based and crowdsourcing approaches. The findings from this study could lead to developments of game mechanisms that are more scalable and cost effective at saliency elicitation than existing paid crowdsourcing approaches.

1 Introduction

Successful product designs commonly acquire shape features that can effectively convey designers' intents. For example, "character lines" are considered as a core design element to establish product identities (e.g., sportiness and luxuriness of cars), and geometric symbols are often created to explain object affordances (e.g., to tell how objects should be interacted with). The mapping from visual information expressed through shapes to design intents is naturally learned by human beings. In fact, it is proven in early studies [1,2] that human beings acquire a superior encoding mechanism to turn high-dimensional sensory information into perceptions. Understanding how shape features contribute to the perception (e.g., the identification of objects, functionalities or semantics) is valuable for various shape design applications, including shape matching [3,4], synthesis [5], texture mapping [5], deformation transfer [5], shape approximation [6,7], and viewpoint selection [6,8], among many others. Further, collecting and learning the map between shape features and perceptions could allow designers to efficiently recreate intended functionalities and semantics for new shapes, or differentiate the identity of a new product from existing ones.

These applications motivated us to consider the fundamental challenge of how such knowledge on shape features shall be collected and validated. To start, we assume that each shape acquires a latent *saliency map*, i.e., a scalar function defined on the meshed geometry, where the key shape features have high saliency values. Therefore, extracting the saliency map will be sufficient for identifying the key features of a shape. The existence of such maps is commonly assumed in saliency extraction studies [4,6–9], yet consensus is to be reached on two critical questions: (Q1) What is the quantitative definition of *saliency*? And (Q2) how can

*Address all correspondence to this author.

we measure the goodness of a given saliency map? To answer Q1, two lines of investigations have been conducted: The first approach hypothesizes that saliency can be measured through shape properties such as local extremities [8], curvature change [6], shape similarity [3], and graph Laplacian [7,9]. The second is to directly elicit saliency maps through crowdsourcing [10], by assuming that human beings share common understanding of salient features. For Q2, validation of computational approaches is often through demonstrating the utility of saliency maps in particular applications, such as viewpoint selection [6] and mesh segmentation [7], or through visual comparison with the crowdsourced answers [7, 8] As a relevant note, a recent research suggested that human visual system may use salient features different from those computed from network models by showing a significant discrepancy between human and computer recognition accuracy along various sizes and resolutions of visual inputs [11].

This paper discusses potential solutions to two major drawbacks among existing methods. First, there lacks a universal way to quantitatively compare the goodness of saliency maps derived from different approaches. Secondly, while existing methods are shown to be useful for certain applications, there is a lack of causality between the heuristics used for the derivation of saliency maps (e.g., change in curvature) and the claimed superior performance (e.g., in viewpoint selection). The proposed solutions are based on the following saliency definition: We consider salient features as a subset of faces from a shape, and assume that, among all subsets of the same size, revealing the salient features will achieve maximal probability of correct recognition of the shape by human beings. Based on this definition, we can measure the goodness of a saliency map by testing the recognition accuracy of the shape through crowdsourcing, for varying amounts of occlusions determined by the map. We then propose a crowdsourcing game where the mechanism is designed to directly elicit saliency maps following the proposed definition.

As a benchmark, we compare three sets of saliency maps (the proposed one, the Schelling point approach [10] and the computational approach [6]) on a subset of shapes from the Princeton database [12] using the viewpoint selection test and the proposed accuracy test. Results are intriguing: First, the proposed saliency maps achieve the most robust performance in the viewpoint selection test (where a view angle is picked to maximize the visible saliency), suggesting that the game mechanism successfully elicits salient features from shapes. Secondly, the three approaches, while being significantly different in saliency maps, are not statistically different in accuracy tests, indicating that various saliency maps may exist according to the proposed saliency definition.

The rest of the paper is arranged as follows: Section 2 reviews related methods used in both crowdsourcing and mesh saliency. In Section 3, we give a new definition to shape saliency and propose an online crowdsourcing game for saliency elicitation. We show our result in Section 4 with discussions. Section five concludes this work.

2 Related Work

In this section, we will review existing studies in shape saliency and crowdsourcing that are related to the presented work. Besides, some applications of saliency will also be discussed.

2.1 Shape saliency

Although there exists a wide range of methods for saliency elicitation, few formal definitions on shape saliency has been given. Early definition of saliency was based on the information theory [1], where salient portions of objects are regarded as such that can be used to infer other portions. Another saliency definition in a 2D context is given by [13] based on contrasty: It defines saliency as those regions that stand out relative to their neighboring parts. Our saliency definition of 3D shapes is more related to that from [1].

Research on 3D shape saliency is built upon studies on 2D images. There is a vast literature on extracting features (or filters) from 2D images in the field of computer vision and machine learning. Feature extraction mechanisms can be found from standard libraries such as OpenCV. The Difference of Gaussian (DoG) and SIFT filters are among the prevalent filters that have been transferred from 2D to 3D applications [6,7,9]. In addition to conventional filters, Convolutional Deep Belief Networks (CDBN) and Convolutional Neural Networks (CNN) have also been considered as a way to extract salient[1] image features [14, 15].

Dedicated saliency extraction methods for 3D objects are often curvature based. A notable example of this kind was proposed in [6], which introduced the idea of scale-invariance features for saliency extraction: Firstly, Gaussian-weighted average of the mean curvature using different variance parameter is used to capture the saliency at different scale; then a non-linear suppression operator [16] is adopted to combine those saliency maps at different scale together to eliminate those with high curvature but repeated patterns. While curvature-based methods are widely adopted, there are cases where curvature information alone is not sufficient to capture perceptually important regions on a 3D shape. By tracking eye movements, it is found that extremities also play an important role in human recognition [17]. This finding motivated researchers to consider extremities as well in computing regional saliency [8]. A similar idea was also applied to a crowdsourcing approach for saliency elicitation: In [10], participants were told to select points they believe that others will also choose, which turned out to be extremities of shapes as well.

More recently, studies pointed out the necessity to consider high-level global saliency besides the regional ones. Log-Laplacian spectrum based method were proposed to this end: Song et al. [9] showed that irregular frequencies in the spectrum of mesh Laplacian is related to global geometry saliency. And spectral Irregularity Diffusion was introduced to preserve both local and global saliency information. However, the challenge of

[1] In the sense that such features contribute to the explanation of image labels

balancing global and local saliency was not resolved. Song et al. improved their methods in [7], however, their results could still be improved for some tested shapes.

2.2 Crowdsourcing and game with a purpose

Crowdsourcing utilizes a large anonymous crowd to complete tasks that are usually hard for computers while intuitive for human beings, e.g., recognition of 2D images and 3D shapes. Since the seminal work on CAPTCHAs[2] [18], crowdsourcing has being successfully applied to a large variety of application domains: 500,000 hours of human efforts per day were achieved on digitalizing books through CAPTCHAs [19]; 14 million images were labeled through ImageNet, revolutionizing the practice of image recognition [20].

One particular type of crowsourcing related to this study is called "game with a purpose" or "gamification", which is to harness human computation through specially designed games. It is estimated that human beings spend 3 billion hours on playing online games every week [21]. When fully utilized, the "wasted" neurological cycles could become an invaluable computation source. This is realized by a variety of attempts: EteRNA [22] connected more than 37,000 enthusiasts to solve RNA design puzzles. EyeWire [23] is another example that gathered 100,000 players to identify neural connections within 3D brain scans. Foldit [24] has attracted 300,000 players (as of 2013) to optimal protein structures. EcoRacer [25] is a recent game for crowdsourcing optimal design and control solutions for vehicle driving.

Crowdsourcing for shape information can be tracked back to as early as the fifties [1]. Recent developments include "Label me", a web-based image annotation tool, which was used to label the identity of objects and where they occur in images [26]. Similarly, ESP game [27] collects semantic descriptions of images by asking a pair of players to come up with the same image descriptions to score. As a follow up of ESP game, the Peekaboom game [28] aims at helping to develop computational algorithms for locating objects in images. In this game, one player is asked to reveal parts of an image related to a given label for another player to guess. Millions of data points has been collected in this way so far. Our proposed game is directly inspired by Peekaboom.

Crowdsourcing 3D saliency is first practiced in [10]. In this study, the authors resort to the idea of "Schelling points", i.e., a set of shape vertices that human participants believe that other participants will also choose. This data is aggregated through an AMT task where participants are told to be rewarded more if vertices they choose overlaps with the majority. This idea is inspired by segregated Nash equilibrium, in which people tend to make the same choices without any communication or feedback. The data from [10] has been used as a ground truth for benchmarking computational approaches.

[2] CAPTCHA stands for "Completely Automated Public Turing test to tell Computers and Humans Apart".

2.3 Application of saliency

Mesh simplification is a classic application of mesh saliency. Better perceptive quality could be achieved, by taking saliency information into consideration during the mesh simplification process [6, 7]. Finding the best viewpoint for an object is also of interest in computer vision and shape processing applications. A good viewpoint could present the geometry in a most informative way. In [6, 8], the authors show that their approaches based on saliency data achieve better viewpoint selection than a curvature-based approach.

Pioneered in [3], assembly based modeling has the potential to greatly reduce the storage of the shape database and facilitate shape synthesis. The key idea is to assemble new models from existing shape component database. It is observed that models of the same class would share similar features, and shuffling interchangeable components of those models could lead to new designs. The very first step towards assembly based modeling is again the extraction of shape features.

In addition, the knowledge on shape saliency could enable unprecedented applications beyond the aforementioned saliency-aware shape editing tools. For example, saliency maps on car bodies could reveal the relationship between shape features and the semantics of the designs (by statistically learning a lower-dimensional embedding of car styles through the revealed features and its mapping to categorical semantic labels). A large scale saliency database would provide a more holistic view of the effects on semantics through shape features than individual designers, and thus may enhance the quality of automated design suggestions (e.g., tuning of character lines) from current approaches where features are manually defined by designers (or researchers) [29, 30]. In addition, a large-scale saliency database could enable automated validation of new product shapes. Conceptually, this can be done by learning a mapping from the geometric representations of shape features to functionality labels, e.g., a handle for grasp often acquires draped surface areas for increased friction. For a given shape design with a "grasp" functionality (this can be inferred from the product identity), the saliency knowledge can be used to evaluate whether the shape can fulfill the design intent or not. Lastly, saliency data could also be used to automate novel shape designs that are mostly differentiated from existing ones from the database, in order to improve brand recognition.

3 Method

In this section we propose a definition of saliency and a crowdsourcing game mechanism that aims to directly extract saliency maps to maximize the recognition accuracy of shapes.

3.1 Saliency definition

Our definition of a saliency map is rooted from [1], which states that saliency represents the predictability of an object based

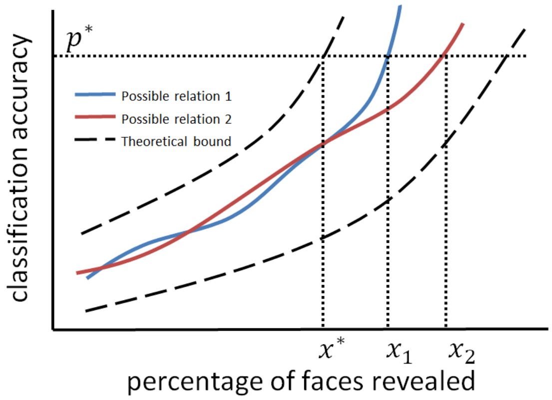

FIGURE 1. Illustration of the saliency definition. The red and blue lines represent two hypothetical mappings from the number of revealed faces to the probability of correct recognition, given different estimates of the latent saliency map. The dashed lines represent hypothetical upper and lower bounds of all possible mappings. (color online)

on some of its portions. We start by assuming the existence of a latent saliency map $s(x) : \mathcal{X} \to \mathbb{R}$ defined on a finite set of faces $\mathcal{X}$. Let $\{x\}_N$ be a shape partially revealed at N percentage. The overall saliency S for $\{x\}_N$ is assumed to be a function: $S_N = f\big(\{x\}_N\big)$. And the probability of correctly recognizing the shape conditioned on $\{x\}_N$ is assumed to be another function: $p(S_N) : \mathbb{R} \to [0,1]$. The salient features $\{x\}^*$ of the shape given N is defined as: $\{x\}^*_N = \text{argmin}_{\{x\}_N \subseteq \mathcal{X}} p(S_N)$ i.e., the set of faces that is most recognizable. We use Fig.1 to illustrate the idea: Here, the solid blue and red curves represent two possible mappings from the percentage of revealed faces N to the probability of correct recognition p, under different estimates of the latent saliency map. The dashed lines are the hypothetical upper and lower bounds of all mappings. For a given probability threshold p^*, the set $\{x\}^*$ derived from the upper bound represents the true salient features. The task of identifying salient features can thus be sufficiently addressed by finding a good estimation $\hat{s}$ of the saliency map. In addition, under the assumption that human beings have shared shape recognition capability, the quality of $\hat{s}$ can be measured by the accuracy p for a variety of occlusion rates given by N.

3.2 Impressionist

Based on the saliency definition, we developed a game called "Impressionist"[3] to elicit saliency maps. Our game can be considered as a 3D extension of the Peekaboom game by Von Ahn [28], which was used for object recognition in 2D. It is also related to a recent crowdsourcing experiment [10], which we will discuss shortly. Our game involves two players, who are randomly paired to prevent direct communication. In each round, one player is shown a shape and asked to reveal faces of the shape for the other player who can then make a guess on the shape. Wrong guesses will be fed back to the first player. Once a correct guess is made, the two switch roles. At the beginning of the game, the players are instructed to reveal as few faces as possible. And after 20 successful guesses as a team, the average face revealing rate is shown to the players along with their ranking among all existing players. Since players are incentivized to compete on releasing less information and making correct guesses, it is hypothesized that they will resort to only revealing the most salient regions of shapes, in which case the cumulative saliency maps derived from the game will lead to the best recognition rate.

It is worth noting that implementation details could influence the effectiveness of this game a lot. In an early test, the correct labels of shapes are not shown to the player who reveals the shape. We noted that this may lead to an inferior face revealing strategy where certain salient features indicated by the correct answer could be ignored. For example, to identify a "duck", one would need to see its bill. However, without knowing that the correct answer is a duck, the player could focus on revealing its wings, which may result in the wrong answer "bird". The combination of the correct labels and the feedback of wrong answers allows the player to adjust his or her revealing strategy during the game, and thus improves the quality of the resultant saliency maps. We have also tested several competition mechanisms to make the game more appealing. In the early test, we ask players to compete on the number of shapes recognized within a limited amount of time. This approach made the game intensive, yet players are not given enough time to plan their face revealing strategy, nor could we collect enough data from the game since each game ends quickly. These findings led to the competition on face revealing rate, which encourage players to spend time on wisely plan their strategies and making guesses on the shapes. While limited observations have been made on players in a lab environment, we found that this second competition mechanism is more preferred by these players.

The game is developed on a Node.js server and uses Socket.IO for player-to-player communications. Real-time 3D rendering is achieved through WebGL and Three.js. A Postgress database is used for data storage. The game interface is designed to be compatible with mobile devices. To allow comparison with existing studies, we use a set of 23 shapes from a standard Princeton geometry database [12] that have been used in a previous crowdsourcing saliency study [10]. The tested shapes are of various complexity and categories, with the number of faces ranging from 1400 to 59640.

4 Experiment and results

This section discusses the experiment settings and the results. We start by presenting the extracted shape saliency from Impressionist players and present results from a view angle test

[3] http://impressionist.herokuapp.com/

to show the differences between our approach and two existing ones [6, 10]. We then compare saliency maps derived from the Impressionist game with those from [10] and [6] with respect to shape recognition accuracies.

4.1 Saliency extraction

Cumulative saliency maps are collected on 23 shapes from the standard Princeton database in a lab setting: A total of 34 voluntary participants were invited to the authors' lab and were paired to play the Impressionist game. The players are given instructions on that (1) they are competing for the minimum revealment of faces and (2) they are not allowed to communicate other than through the game. On average, every object received 24.3 guesses, and 15.4 of them are correct. To create saliency maps from game plays, we adopt a random work approach [31] to diffuse the accumulated number of selection across the shape. This post-processing is similar to the Schelling distribution function applied in [10]. Saliency maps of different shapes are visualized as heat maps, as shown in Fig. 2. These maps are plotted in view angles that maximize the sum of all visible saliency, as proposed in [6]. Followed by the same procedure, the saliency maps derived from two existing studies [6, 10] are plotted for comparison in Figs.3 and 4, respectively.

From the comparison, the advantages of our method can be seen: (1) First, our method robustly captures regional saliency. It is noticed that saliency maps from [10] almost only capture the tips of shapes. This is due to the inherent flaw that only extremities are meant to be captured as Schelling points. However, salient lines and regions often exist in shapes. Unlike Schelling points that are only concentrated on a small number of local extremes, Impressionist reveals varying saliency patterns across shapes. For example, the nose of a pig shape has high saliency concentration while a giraffe obtains across its neck smeared saliency. The "woman" shape has high saliency on chest from our result while the face and joints are highlighted by [10]. For some shapes, the saliency obtained from curvature method is similar to Impressionist, e.g., the pair of glasses and the plane. Nonetheless, we also observe such shape whose high curvature areas are not salient at all, e.g., duck and bird, resulting in unfavorable view angles by [6]. This leads to the second advantage of our method: Compared with existing methods, Impressionist data consistently produce reasonable view angles across shapes.

4.2 Classification accuracy

We now compare the goodness of these three saliency elicitation methods based on the proposed definition of saliency. Based on the collected saliency maps, a survey[4] was conducted through Amazon Mechanical Turk. In each survey, 10 occluded shapes from the aforementioned 23 shapes were shown to the participant in a sequence, who is then asked to recognize the shapes. For each

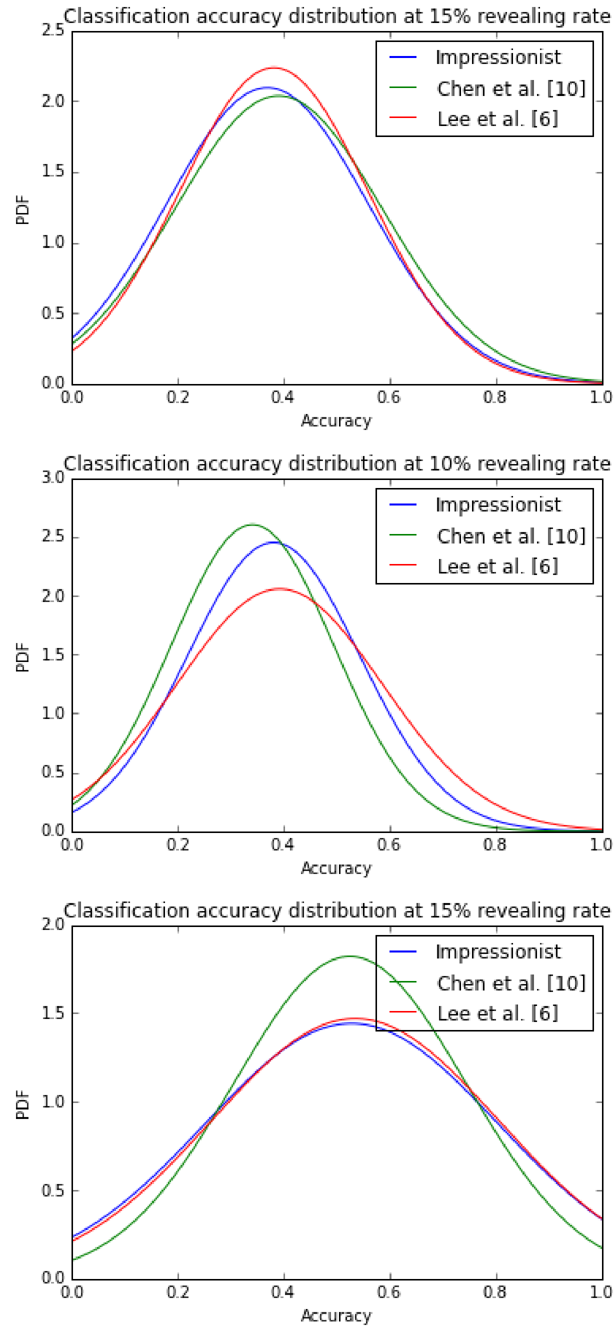

FIGURE 5. A comparison on recognition accuracies of the three methods, with 5%, 10% and 15% of faces revealed to the participants.

shape, a face-revealing rate of 5%, 10% and 15%, as well as one of the three saliency maps, were uniformly chosen. A total number of 7,035 guesses were collected. By assuming that the recognition accuracies across shapes are indepently drawn from a normal distribution, t-tests are conducted to test the null hypotheses that

[4] http://impressionist.herokuapp.com/cmp

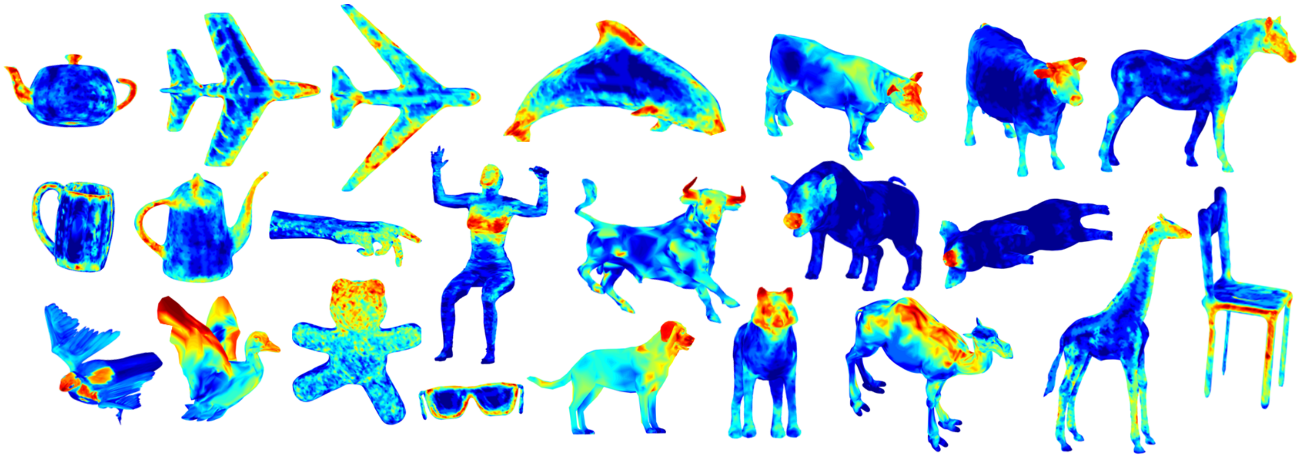

FIGURE 2. Saliency maps derived from Impressionist, shown in best view angles.

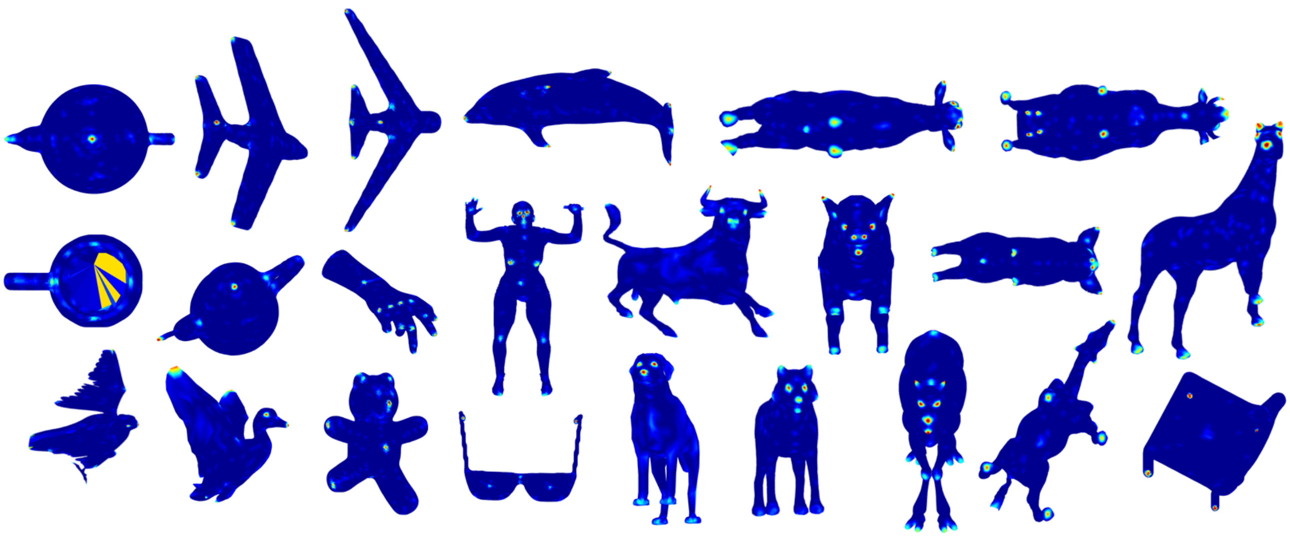

FIGURE 3. Saliency maps derived from [10], shown in best view angles.

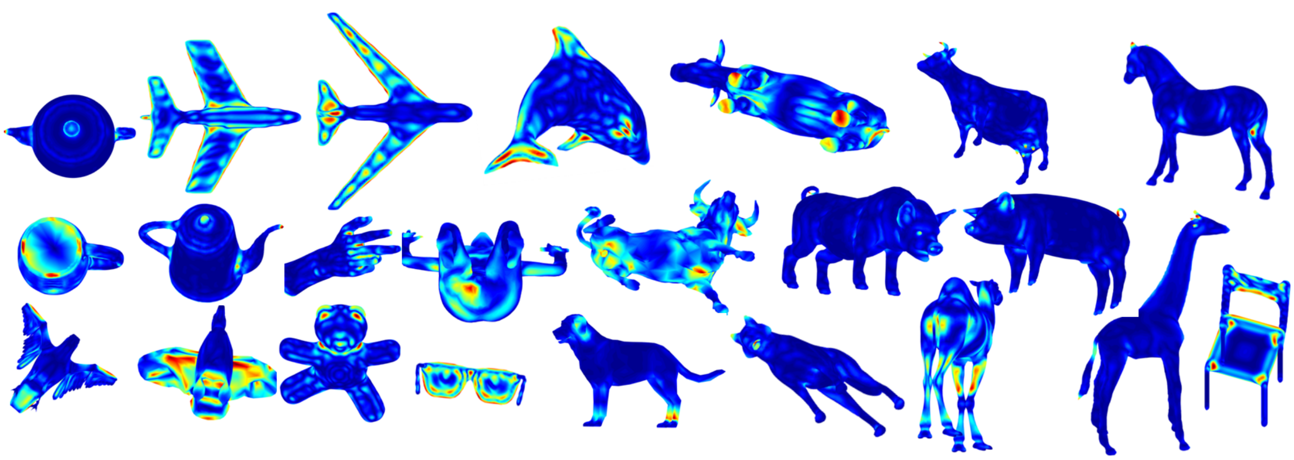

FIGURE 4. Saliency maps derived from [6], shown in best view angles.

each pair of methods has the same recognition accuracy. Results of Impressionist were compared to [10] and [6], as shown in Fig.5. While the differences in saliency maps are significant, it is interesting to see that the resultant recognition accuracies from these three methods are statistically non-differentiable, with the only exception for the comparison between Impressionist and [10], at 10% of revealing rate (with a p-value of 0.02). This result indicates that drastically different saliency elicitation methods could provide similar recognition accuracies, and that human beings can adapt to both global and local salient features when they infer shape identities.

4.3 Limitations and potential solutions

A few limitations have been identified through the development of Impressionist. First, while the user interface and the competition mechanism are carefully designed, the popularity of the game is still limited. Significant effort is needed to enhance the entertainment elements for the game and promote the game to the general public, in order to fulfill its potential at saliency elicitation. Besides, as all two player games, the current implementation needs to avoid player playing against himself. In the future, an IP check or an account system will be implemented to avoid this issues. We also note that players, after several plays, could learn from the shape database and achieve more efficient guesses, since probability masses are now concentrated on shapes that they have seen in previous rounds. While this effect may not negatively affect the quality of the cumulative saliency maps, it could discourage players from further contributing their knowledge. This issue will be alleviated by increasing the size of the shape database. Lastly, shapes often acquire multiple equivalent labels and labels from various levels of refinement. Each label may imply a different saliency map. Future work will investigate how changes in descriptions (labels) will infect saliency map for the same shape.

It is also worth noting that the proposed game mechanism produces more interesting data than saliency knowledge. For example, we noticed that players produce wrong guesses quite frequently during the game. These wrong guesses carry valuable information about the shapes. For example, they can be used to measure the similarity among shape labels; and the faces revealed between a wrong and a correct guess could be used to tell apart the two similar shapes. There is another important advantage of our method compared with existing computational method. Other information on geometry like texture could also be considered in Impressionist directly, no further changes are needed at all. The investigation of texture on saliency would be carried out in the future.

5 Conclusions

Documenting salient shape features and understanding how they convey the funtionalities and semantics of products will enhance design practices and promote salient-aware design tools. This study is thus motivated to investigate the fundamental challenge in elicitation and validation of saliency maps on shapes. Our contributions are twofold. First, we proposed to measure the goodness of saliency maps by shape recognition accuracy under occlusions determined by the maps. Secondly, we designed and tested a game mechanism, Impressionist, that could extract shape saliency from human players.

We conducted experiments to show that: (1) Saliency maps derived from Impressionist are more robust at viewpoint selection, while various less meaningful viewpoints were produced by both the Schelling point method [10] and the curvature method [6]. This result suggested that the proposed game mechanism fulfills the purpose of eliciting salient shape features. (2) Saliency maps derived from curvature change [6], Schelling points [10] and the proposed game achieved statistically similar recognition accuracies and are thus equally valid from this perspective. The fact that all three methods achieve non-differentiable recognition accuracies across the tested shapes suggests that human beings can recognized shapes equally effectively through either global or local features. This conjecture, however, requires further investigation.

To summarize, the proposed gamification approach for saliency elicitation has compatible performance in shape recognition to existing crowdsourcing and computational approaches and is more reliable in viewpoint selection. In addition, the use of a game would potentially allow more cost-effective elicitation at a scale beyond existing paid crowdsourcing models. Future research will investigate competition and collaboration mechanisms to further improve the quality of the elicitation and self-validation by players, as well as the appeal of the game. Developments and findings from this paper will lead to novel ways of documenting design intents and corresponding salient shape feature at an unprecedented scale.

Acknowledgement

This work has been supported by the National Science Foundation under Grant No. CMMI-1266184 and the startup funding from Arizona State University. These supports are gratefully acknowledged. We would also like to thank Dr. Hanghang Tang for his generous support, and all crowd participants for their contribution.

REFERENCES

[1] Attneave, F., 1954. "Some informational aspects of visual perception.". *Psychological review,* ***61***(3), pp. 183–193.
[2] Shannon, C. E., 1951. "Prediction and entropy of printed english". *Bell system technical journal,* ***30***(1), pp. 50–64.
[3] Funkhouser, T., Kazhdan, M., Shilane, P., Min, P., Kiefer, W., Tal, A., Rusinkiewicz, S., and Dobkin, D., 2004. "Mod-

eling by example". *ACM Transactions on Graphics,* ***23***(3), p. 652.
[4] Shilane, P., and Funkhouser, T., 2007. "Distinctive regions of 3D surfaces". *ACM Transactions on Graphics,* ***26***(2), pp. 7–es.
[5] Sorkine, O., Cohen-Or, D., Lipman, Y., Alexa, M., Rössl, C., and Seidel, H.-P., 2004. "Laplacian Surface Editing". *Eurographics Symposium on Geometry Processing*, pp. 175–184.
[6] Lee, C. H., Varshney, A., and Jacobs, D. W., 2005. "Mesh saliency". *ACM Transactions on Graphics,* ***24***(3), p. 659.
[7] Song, R., Liu, Y., Martin, R. R., and Rosin, P. L., 2014. "Mesh saliency via spectral processing". *ACM Trans. Graph.,* ***33***(1), Feb., pp. 6:1–6:17.
[8] Leifman, G., Shtrom, E., and Tal, A., 2012. "Surface regions of interest for viewpoint selection". *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 414–421.
[9] Song, R., Liu, Y., Martin, R. R., and Rosin, P. L., 2013. "3D point of interest detection via spectral irregularity diffusion". *User Modeling and User-Adapted Interaction,* ***29***(6-8), pp. 695–705.
[10] Chen, X., Saparov, A., Pang, B., and Funkhouser, T., 2012. "Schelling points on 3D surface meshes". *ACM Transactions on Graphics,* ***31***(4), pp. 1–12.
[11] Ullman, S., Assif, L., Fetaya, E., and Harari, D., 2016. "Atoms of recognition in human and computer vision". *Proceedings of the National Academy of Sciences,* ***113***(10), pp. 2744–2749.
[12] Shilane, P., Min, P., Kazhdan, M., and Funkhouser, T., 2004. "The princeton shape benchmark". In Shape modeling applications, 2004. Proceedings, IEEE, pp. 167–178.
[13] Borji, A., and Itti, L., 2013. "State-of-the-art in visual attention modelling". *IEEE Transactions on Pattern Analysis and Machine Intelligence,* ***35***(1), pp. 185–207.
[14] Lee, H., Grosse, R., Ranganath, R., and Ng, A. Y., 2009. "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations". *Proceedings of the 26th Annual International Conference on Machine Learning ICML 09,* ***2008***, pp. 1–8.
[15] Krizhevsky, A., Sutskever, I., and Hinton, G. E., 2012. "ImageNet classification with deep convolutional neural networks". In Advances in neural information processing systems, pp. 1097–1105.
[16] Itti, L., Koch, C., and Niebur, E., 1998. "A model of saliency-based visual attention for rapid scene analysis". *IEEE Transactions on Pattern Analysis and Machine Intelligence,* ***20***(11), pp. 1254–1259.
[17] Kim, Y., Varshney, A., Jacobs, D. W., and Guimbretière, F., 2010. "Mesh saliency and human eye fixations". *ACM Transactions on Applied Perception,* ***7***(2), pp. 1–13.
[18] Von Ahn, L., Blum, M., Hopper, N. J., and Langford, J., 2003. "CAPTCHA: Using Hard AI Problems for Security". *Advances in Cryptology - EUROCRYPT 2003,* ***2656***(Lecture Notes in Computer Science), pp. pp 294–311.
[19] von Ahn, L., 2009. "Human computation". *Proceedings of the 46th Annual Design Automation Conference - DAC '09*, pp. 418–419.
[20] Deng, J. D. J., Dong, W. D. W., Socher, R., Li, L.-J. L. L.-J., Li, K. L. K., and Fei-Fei, L. F.-F. L., 2009. "ImageNet: A large-scale hierarchical image database". *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2–9.
[21] Schrope, M., 2013. "Solving tough problems with games". *Pnas,* ***110***(18), pp. 7104–7106.
[22] Lee, J., Kladwang, W., Lee, M., Cantu, D., Azizyan, M., Kim, H., Limpaecher, A., Yoon, S., Treuille, A., and Das, R., 2014. "RNA design rules from a massive open laboratory.". *Proc Natl Acad Sci U S A,* ***111***(6), pp. 2122–7.
[23] Kim, J. S., Greene, M. J., Zlateski, A., Lee, K., Richardson, M., Turaga, S. C., Purcaro, M., Balkam, M., Robinson, A., Behabadi, B. F., et al., 2014. "Space-time wiring specificity supports direction selectivity in the retina". *Nature,* ***509***(7500), p. 331.
[24] Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popović, Z., et al., 2010. "Predicting protein structures with a multiplayer online game". *Nature,* ***466***(7307), pp. 756–760.
[25] Ren, Y., Bayrak, A., and Papalambros, P., 2015. "ecoRacer: Game-Based Optimal Electric Vehicle Design and Driver Control Using Human Players". *ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*.
[26] Aydemir, A., and Sjoo, e. a., 2011. "Perceptual Strategies for Purposive Vision". *International Journal of Computer Vision,* ***77***(2), pp. 3515–3522.
[27] von Ahn, L., and Dabbish, L., 2004. "Labeling images with a computer game". *ACM Conference on Human Factors in Computing Systems,* ***6***(1), pp. 319 – 326.
[28] von Ahn, L., Liu, R., and Blum, M., 2006. "Peekaboom". In Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI '06, Vol. pages, ACM Press, p. 55.
[29] Ren, Y., and Papalambros, P. Y., 2011. "A design preference elicitation query as an optimization process". *Journal of Mechanical Design,* ***133***(11), p. 111004.
[30] Yumer, M. E., Chaudhuri, S., Hodgins, J. K., and Kara, L. B., 2015. "Semantic shape editing using deformation handles". *ACM Transactions on Graphics (TOG),* ***34***(4), p. 86.
[31] Tong, H., Faloutsos, C., and Pan, J. Y., 2006. "Fast random walk with restart and its applications". *Proceedings - IEEE International Conference on Data Mining, ICDM*, pp. 613–622.