



HAL
open science

EVALUATION OF THE BUBBLE VIEW METAPHOR FOR THE CROWDSOURCING STUDY OF VISUAL ATTENTION DEPLOYMENT IN TONE-MAPPED IMAGES

Waqas Ellahi, Toinon Vigier, Patrick Le Callet

► **To cite this version:**

Waqas Ellahi, Toinon Vigier, Patrick Le Callet. EVALUATION OF THE BUBBLE VIEW METAPHOR FOR THE CROWDSOURCING STUDY OF VISUAL ATTENTION DEPLOYMENT IN TONE-MAPPED IMAGES. 9th European Workshop on Visual Information Processing (EUVIP), Jun 2021, Paris (virtual), France. 10.1109/EUVIP50544.2021.9483985 . hal-03373606

HAL Id: hal-03373606

<https://hal.science/hal-03373606>

Submitted on 11 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

EVALUATION OF THE BUBBLE VIEW METAPHOR FOR THE CROWDSOURCING STUDY OF VISUAL ATTENTION DEPLOYMENT IN TONE-MAPPED IMAGES

Waqas Ellahi, Toinon Vigier, Patrick Le Callet

LS2N UMR CNRS 6003, Université de Nantes, Nantes, France
firstname.name@univ-nantes.fr

ABSTRACT

Attention is an important attribute of human vision for study of user's quality of experience (QoE). The attention information collection from eye tracking is impossible in the current scenario of Covid-19. Different mouse metaphors have been proposed to study visual attention without eye tracking equipment. These methods have shown promising results on different types of images (visualizations, natural images and websites) with well-identified regions of interest. However, they have not been precisely tested for QoE applications, where natural images are processed with different algorithms (compression, tone-mapping, etc.) and visual content can induce more exploratory behavior. This paper studies and compares different configurations of bubble view metaphors for the study of visual attention in tone-mapped images.

Index Terms— eye movements, visual attention, crowdsourcing, eye tracking, Bubble view

1. INTRODUCTION AND RELATED WORKS

The study of visual attention (VA) through eye tracking (ET) data is valuable in various domains, such as neuroscience, quality of experience (QoE), etc. In the QoE domain, VA can be seen as a proxy for the perceived experience and it has been used for numerous applications as image and video compression [1], image memorability [2]. However, the collection of gaze data has been difficult at large scale, as it needs in-lab ET using dedicated hardware. Involving participants and running experiments in sequence takes plenty of time due to limited resources. Meanwhile, large amounts of participants' data can be collected with the help of online crowdsourcing. Currently, the importance of crowdsourcing as technique is increased further due the widespread of COVID-19. In such circumstances, it is very difficult to enroll participants for the lab experiments. Initially, webcam-based eye tracking was considered as an alternative for crowdsourcing [3, 4]. However, it requires such as participant pose and specific lighting conditions. These conditions are difficult to enforce for an online experiment. Additionally, to access participant's camera raises privacy and ethical concerns. These issues have motivated the research community to adopt alternate solutions for capturing attention data without any ET device.

In this context, several metaphors have already been proposed to collect attention data without ET. Recently, the TurkEyes project made available online four of these methods and published a detailed benchmark of these methods on different contents (natural images, graphic designs, and data visualizations)[5]. **Import Annots** [6] asks the participants to paint over interesting regions. The binary masks are then used to detect salient regions. This method is more suitable for informative content and not ideal for natural images. **Zoom Map** [5] uses zoomable viewports and treat them as areas

of interest from which it is possible to build heat maps. This method is more suitable for multi-scale content and gives a coarse approximation of VA. **Code Charts** [7] ask participants to specify on a grid where they gazed on an image. This method gives well-correlated results with eye data but it allows to acquire only one fixation per observer, thus it cannot be used to study individual attention deployment and scanpaths. **Bubble View** (BV) [8] consists in blurring the content except the region around the mouse cursor after clicking. This method can be seen as a good metaphor of eye movements and shows good correlation with eye data.

Some adaptations of the BV metaphor without clicks have been proposed, mainly for natural images [9, 10]. Indeed, the task of clicking misses the capture of bottom-up and unconscious visual explorations. In the SALICON dataset [11], the authors used an adaptive blur closer to the human visual system but complex in terms of computations leading to possible problems of performance for crowdsourcing studies.

In the context of QoE, it is argued that the chosen mouse-based metaphor to study VA should: 1) acquire bottom-up and top-down processes; 2) enable the possibility to work on idiosyncratic aspects by studying individual VA deployment; 3) enable to study not only visual saliency but also scanpaths in order to consider the temporal aspect of VA; 4) be applied on degraded images with various artifacts from different processing techniques; 5) be applied on exploratory images i.e., images without strong salient objects. Considering the first three constraints, it is concluded that the BV metaphor without clicks can be a good technique to develop crowdsourcing studies of visual attention in QoE image datasets. However, it is necessary to ensure the validity of this method for QoE visual content. Indeed, they often include a large variety of images in terms of spatial information, contrasts, textures and number of regions of interest whereas the BV technique has been tested only on two natural images databases mainly dedicated to salient objects detection (MS COCO [9] and OSIE [10]) which include pristine images with clear objects of interests.

This present work evaluates the BV metaphor in the context of tone-mapped images. Tone mapping consists in the compression of the dynamic range of an image to match the characteristics of the displays, preserving the aesthetic visual quality of the represented scene and the artistic intention. Previous works based on ET experiments showed that tone mapping operators can significantly impact the fixation behavior highlighting the strong effect on contrast in VA [12]. For these reasons, tone mapping is a good representative use case of QoE applications. The paper is structured as follows: Section 2 covers experimental details. Section 3 discusses processing of BV data. The results and discussion are covered in Section 4, whereas the concluding remarks are provided in Section 5.

2. EXPERIMENTS

2.1. Eye-tracking Experiment

The ET dataset was collected by inviting 55 participants in a lab experiment to perform a free task. The 10 source HDR images and 9 different HRC (4 tone mapping operators (TMO) with two different setting parameters and one TMO with single setting) were selected for the experiment [13]. The Tobii Pro Fusion eye tracker of 120 Hz was used for gaze recording. The experiment was designed in Tobii Pro Lab software [14]. The image content was displayed on a 24-inch screen with 1920x1080 resolution and the lighting conditions followed the ITU standards. Each stimulus was presented to participants for 5 seconds followed by a fixation cross mark image shown for 2 seconds between consecutive stimuli. The protocols necessary due to Covid-19 were carefully followed during the experiment and the experiment room was properly sanitized before and after every participant. For this work, only the data collected from 28 participants in the first session is considered in order to reduce the memorization effect.

Table 1: Characteristics of both view experiments.

| Features | ET | Config.1 | Config.2 |
|----------------|-----------|---------------|------------|
| # of images | 90 | 20 | |
| Resolution | 1920x1080 | | |
| Time (sec) | 5 | 20 | 10 |
| Distance | 3 H | - | - |
| # of Observers | 55 | 20 | 21 |
| Ages | 33.4±11.9 | 30.1±10.1 | 29.5±7.3 |
| Gender (M/F) | 20/35 | 15/5 | 19/2 |
| Eyetracker | Tobii Pro | - | - |
| Bubble size | - | 2°(120 px) | 4°(240 px) |
| Blur sigma | - | 0.278°(16 px) | |

2.2. Bubble View Experiment

In this paper, the BV method is applied without clicks in order to consider both bottom-up and top-down perceptual processes. Three different parameters must be defined for running a BV experiment. 1)**Bubble size:** The bubble size matters a lot as it must simulate the human vision (fovea) without creating difficulty in the task with too small bubble size[15]. 2)**Blur sigma:** The blurring threshold must also simulate the human vision (peripheral vision). A trade-off should be found to keep some details in order to elicit attraction to specific regions without giving too much precision to force exploration with the bubble. 3)**Stimuli display time:** The time also has a significant impact as less time forces the participants to see only the most interesting regions whereas additional time can lead to the elaboration of specific strategies to explore the image as a top to bottom scan for instance.

Different configurations have been tested in previous works using the BV metaphor without clicks on natural images. In [8], different bubble sizes and levels of blur were tested from 30 to 50 pixels. Results show a significant effect only of the blur effect. However, these values are not discussed from perceptual point of view, i.e. number of pixels per visual degree based on realistic conditions of crowdsourcing experiments (e.g. viewing distance, screen size and content resolution). In [11], the level of the progressive blur is computed based on 7.5 pixels per visual degree which seems not ecologically valid. The mouse-based saliency maps (SM) are then



Fig. 1: Screen shot of BV experiment

computed with different values of sigma in order to achieve the best performance of the dataset without clear perceptual justification.

It is thus proposed to explore different configurations for BV crowdsourcing experiments, based on perceptual parameters computed from ecological situations. The bubble size is chosen from 2 to 4 degrees to stimulate foveal to near-parafoveal areas. The level of blur is then chosen to represent the maximal eccentricity considering the vertical viewing angle. In order to express these values in pixels, we need to know the experimental setup which is basically not the case in crowdsourcing experiments. Thus, some information about the viewing conditions of crowdsourcers is assumed. For this experiment, a Full HD resolution was required and we assume that the users were equipped with a laptop of 14 inches for a viewing distance of 55 cm, which is in line with ergonomics recommendation. This corresponds to a vertical viewing angle of 18 visual degrees and 60 pixels per degrees. The blur level was defined as the visual acuity for an eccentricity of 18°, i.e., 0.12 Snellen Fraction corresponding to 3.6 cycles per degree, thus a sigma of 0.278° (approximately 16 pixels). It was postulated that exploring an image with a mouse could be longer than naturally with eyes. Moreover, a smaller bubble size can require a longer duration to explore the same sized image. At the end, the experiment was set up for two configurations described in Table 1. As the objective of this experiment was to study BV configurations, the dataset was reduced to two TMOs (Drago1 and Mantiuk1) for the 10 HDR sources. The Prolific platform was used for conducting the experiment [16] and the mouse data of 20 and 21 participants were respectively collected on configuration 1 and 2.

3. DATA PROCESSING

3.1. Description of the mouse data collected during BV experiment

The cursor movement events were recorded for each participant for each source during an experiment. The cursor event is described by cursor position (x, y) on the screen at a time-stamp (ts). The distribution of distance and duration between two cursor events is presented in Fig. 2. A larger distance between two events can be noticed for a bigger bubble size. However, the duration between two events, which corresponds to the time spent at a specific position, is comparable in the two configurations, i.e., around 20 milliseconds which is far lower than a fixation duration. We can assume in BV metaphor that observers process visual information through smooth pursuits following mouse trajectories rather than doing eye fixations. Thus the duration of between event would match the speed of the smooth pursuit (< 30 degrees/sec). This assumption as well as the consequence on visual attention deployment in this kind of experiment should be further studied.

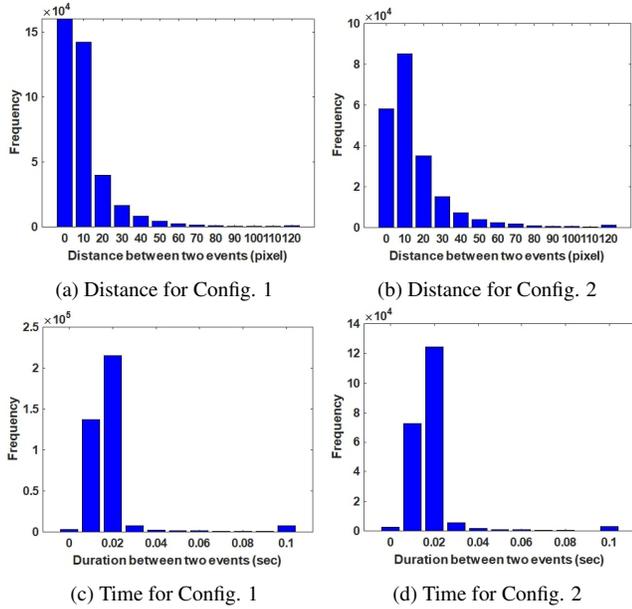


Fig. 2: Distribution of distance and time between two mouse events.

3.2. Outliers detection

The crowdsourcing platform has many advantages including the access to a large population with less expense. However, some challenges also exist, such as trustworthiness of the internet users. The untrustworthy participants may provide erroneous response [17]. Such erroneous responses are useless and may deviate significantly from other users' responses. So, outliers have to be identified and removed in order to achieve a robust evaluation. The outlier or spam detection for crowdsourcing itself is an open research problem. The attention of participation and trusty response are judged by the number of mouse movement events. The number of events is considered as an indicator for outlier identification in this paper.

This method is simple but effective to detect strange behavior in participants. The participants who never (not frequently) move their mouse or people who move the mouse with too high speed without exploring the region of BV are referred to as an outlier. Further improvements of outlier detection is a topic for future studies.

3.3. Time-based saliency map

Fig. 3 shows the mouse events for one specific stimulus in comparison with ET gaze data. It is clearly evident that the plot of mouse events does not exhibit fixation clusters as eye data. This could be explained by the non-fixed sample rate for BV experiment, the absence of "micro-saccades" in mouse-based exploration as well as the possible smooth pursuits of the bubble rather than fixations. It is thus necessary to find a way to consider mouse-based fixation areas.

In this paper, we propose to apply a simple method based on the accumulation of the time spent at one specific pixel position for computing SM. As for eye-tracking based SM, a Gaussian blur with a sigma of one visual degree emulating fovea and a kernel size equal to the bubble size is then applied. To summarize, the BV SM (BVSM) is computed as follows:

$$BVSM = G_{\sigma}^k \left(\sum t(x, y) \right) \quad (1)$$

where $t(x, y)$ represents the time spent at specific position, σ and k are sigma and kernel size of the Gaussian filter (here $\sigma = 1$ visual degree and $k =$ bubble size). The SM computed using this method is presented in Fig. 4

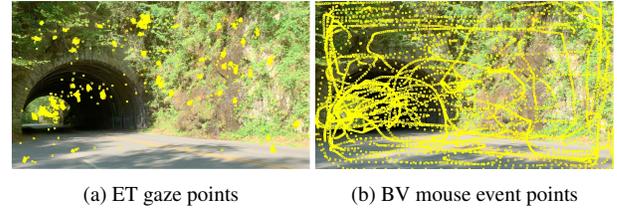


Fig. 3: Plots of collected data on image

This method is different from the one presented in [11] which consists in the re-sampling of the mouse events at a fixed rate as well as the exclusion of the samples with high velocity considered as saccades. One drawback of the re-sampling method is the need to define a threshold from which all the faster points will be removed.

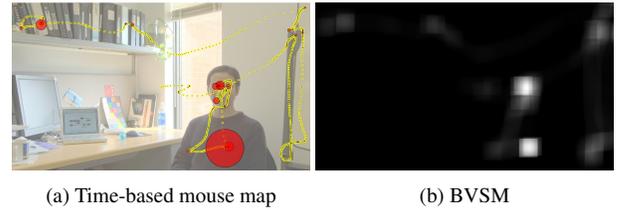


Fig. 4: Time-based mouse map (the size of red circle represents the time spent of each event) and the associated BVSM.

4. RESULTS AND DISCUSSION

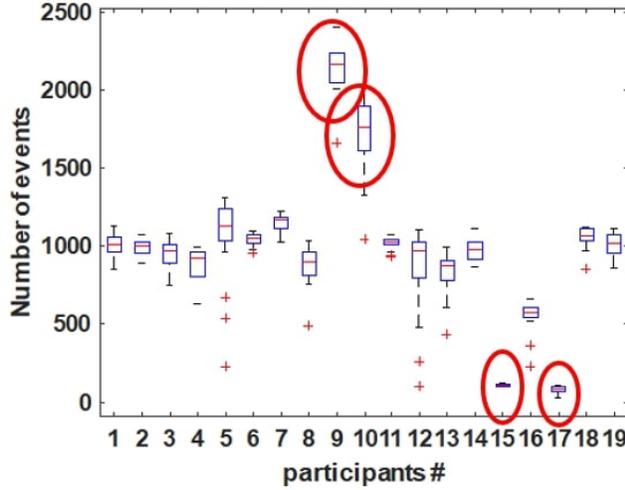
4.1. Outliers

As a first step, one participant is removed for Config.1 due to incomplete response for one stimulus. Then, as explained in 3.2, outliers are removed by considering the number of samples in the response per source. Fig. 5 represents the number of samples per source per participant. In this figure, it can be noticed that some participants moved cursor very fast or very slow and did not properly give attention to the experiment. Therefore, additional 4 participants in Config. 1, and 3 participants in Config. 2 are removed. So, a total of 15 participants data for Config. 1 and 17 participants data for Config 2 are used for further analyses.

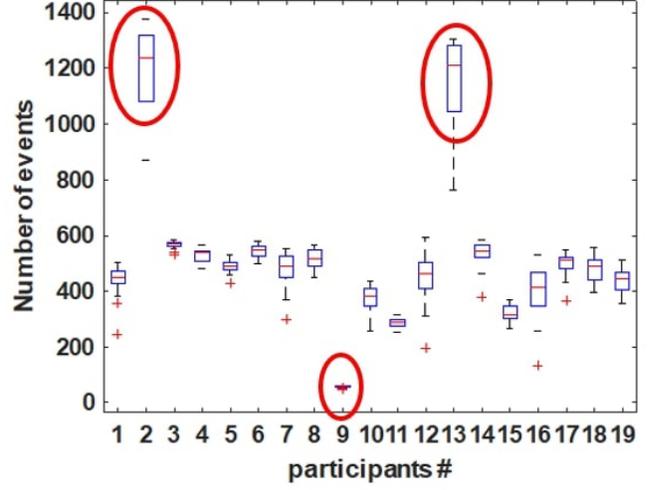
4.2. Similarity of BV method with ET data

In order to assess the similarity of the mouse-based metaphor with ET, the similarity between the BV and ET SM is computed with the Kullback-Liebler Divergence (KLD), correlation (CC) and Normalized Scanpath Similarity (NSS) metrics [1]. As proposed in [5], this comparison is made using participants' baseline scores estimated from ET data. In particular, for CC and KLD, the baseline score is computed by randomly splitting the participants ET data into two groups. The similarity is measured between two SM computed from the corresponding two groups.

For NSS, the baseline score is computed by employing inter-observer congruency (IOC) metric, i.e., the average of individual



(a) Config. 1



(b) Config. 2

Fig. 5: Number of mouse events recorded per participant for the two different configurations.

NSS scores computed between the global ET SM compared with the gaze data of each participant. In our proposed analysis, the baseline score is computed per stimulus since each source contains different visual characteristics which affects the number of regions of interest and inter-observers' congruency. The results of similarity are also computed as a percentage of ET observer consistency by normalizing the scores with the baseline as in [5, 8]

In order to compare the different characteristics of the BV experiment, the similarity scores are computed for each configuration and for different viewing durations. The similarity measures versus the baseline scores are plotted in Fig. 6 and summarized in Table 2 and 3. Except for very few cases, SM between BV and ET saliency maps is lower than the baseline as expected. The plots (Fig. 6) exhibit an important effect of content on BV and ET similarity results, which confirms that mouse-based metaphors need to be evaluated for QoE applications with ad-hoc datasets. Moreover, we cannot see evident correlation between baseline and BV scores. The inter-observer congruence is not sufficient to explain the efficiency of BV from one content to another, and further analyses should be led to better study the impact of visual content features.

Concerning the BV parameters, results show better scores for a viewing duration of 5 seconds in Config. 1, i.e., a bubble size of two visual degrees. For these specific parameters based on perceptual features, BV performance on our tone mapping dataset is comparable to TurkEyes results obtained on specific dataset for the detection of salient objects [5]. These findings suggest that BV can be an adequate mouse-based metaphor for the study of VA for QoE applications. However, further studies on other datasets with different types of content (artefacts, videos, etc.) should confirm these results. Moreover, the viewing duration of 5 seconds should be validated in real conditions.

4.3. Impact of TMOs on VA with BV technique

One objective of this study is also to see if the BV metaphor can exhibit different attentional behavior in respond to different TMOs, as it has been shown in study [18]. In this section, only BVSM obtained with a viewing duration of 5 seconds are used. Fig. 7 shows the participant VA response to sources processed using two differ-

Table 2: Similarity scores mean (\pm standard deviation) between BVSM and ET data across content for Config. 1.

| Metrics | 5s | 8s | 10s | 13s | 16s | 20s |
|------------------|--------------|--------------|--------------|--------------|--------------|--------------|
| NSS \uparrow | 1.4 \pm .5 | 1.1 \pm .3 | 0.9 \pm .2 | 0.8 \pm .2 | 0.7 \pm .2 | 0.7 \pm .2 |
| % | 57 \pm 18 | 44 \pm 12 | 39 \pm 10 | 33 \pm 9 | 30 \pm 9 | 29 \pm 9 |
| CC \uparrow | 0.7 \pm .1 | 0.6 \pm .1 |
| % | 78 \pm 11 | 74 \pm 12 | 72 \pm 12 | 70 \pm 12 | 68 \pm 13 | 68 \pm 14 |
| KLD \downarrow | 1.9 \pm 1 | 2.3 \pm 1 | 2.5 \pm 2 | 2.6 \pm 2 | 2.8 \pm 1 | 2.8 \pm 1 |
| % | 37 \pm 14 | 29 \pm 10 | 27 \pm 8 | 25 \pm 6 | 24 \pm 7 | 24 \pm 8 |

Table 3: Similarity scores mean (\pm standard deviation) between BVSM and ET data across content for Config. 2.

| Metrics | 5s | 6s | 7s | 8s | 9s | 10s |
|------------------|--------------|--------------|--------------|--------------|--------------|--------------|
| NSS \uparrow | 1.0 \pm .3 | 1.0 \pm .2 | 0.9 \pm .2 | 0.9 \pm .2 | 0.8 \pm .2 | 0.8 \pm .2 |
| % | 43 \pm 10 | 40 \pm 10 | 37 \pm 9 | 36 \pm 9 | 34 \pm 9 | 34 \pm 9 |
| CC \uparrow | 0.7 \pm .1 |
| % | 78 \pm 11 | 78 \pm 11 | 77 \pm 11 | 77 \pm 10 | 78 \pm 10 | 78 \pm 9 |
| KLD \downarrow | 2.0 \pm 1 | 2.0 \pm 1 | 2.1 \pm 1 | 2.2 \pm 1 | 2.2 \pm 1 | 2.3 \pm 1 |
| % | 34 \pm 10 | 33 \pm 9 | 32 \pm 9 | 31 \pm 8 | 30 \pm 8 | 30 \pm 8 |

ent TMOs for ET and BV (Config.1) data. Fig. 7 shows that similar to ET data, BV data also reflects change in VA in response to different TMOs. The analysis is further investigated by computing the similarity scores between ET SM for the same source across two different TMOs using NSS, CC and KLD methods. In an analogous way, the scores are also computed for BVSM. Then, the correlation between the similarity measures between TMOs for ET and BV data is computed. A high correlation indicated that the BV metaphor can be used to discriminate TMOs in terms of VA. The obtained scores, presented in Table 4, show a strong correlation between BV and ET based similarity measures between TMOs for Config. 1. These results confirm that BV with a bubble size of 2 $^\circ$ is a good metaphor to study the impact of TMO on VA in a crowdsourcing experiment.

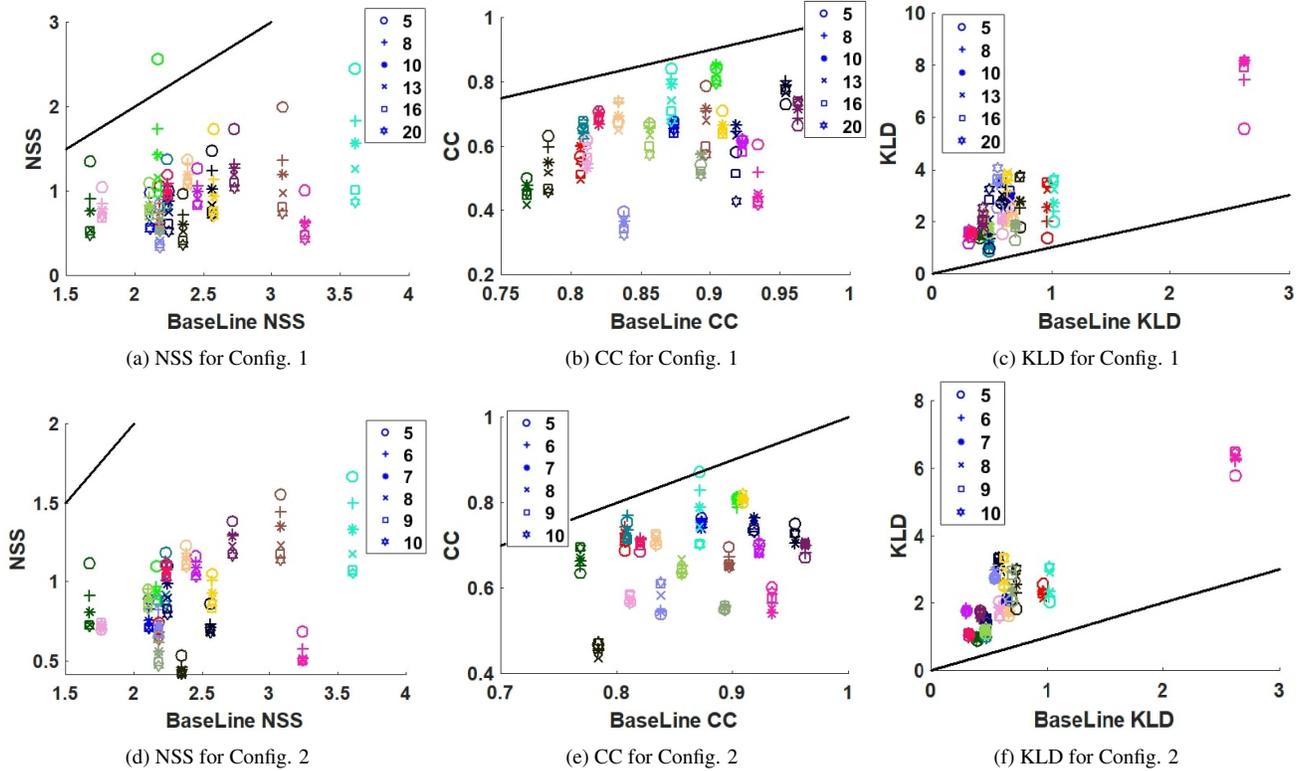


Fig. 6: Plots of metric scores between BV and ET data versus the baseline. The different colors and markers represent stimuli and viewing time respectively. The black line corresponds $X=Y$ line.

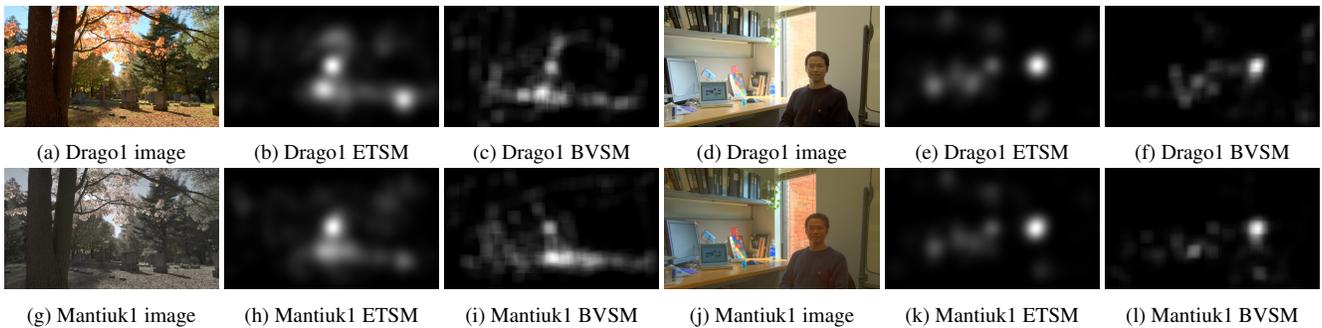


Fig. 7: Effect of tone mapping operators on VA.

Table 4: Pearson (PCC) and Spearman (SORCC) correlation scores between ET and BV based similarity across TMOs

| Correlation | Config. 1 | | | Config. 2 | | |
|-------------|-----------|------|------|-----------|------|------|
| | NSS | CC | KLD | NSS | CC | KLD |
| PCC | 0.79 | 0.70 | 0.27 | 0.49 | 0.56 | 0.20 |
| SORCC | 0.65 | 0.78 | 0.19 | 0.33 | 0.39 | 0.19 |

5. CONCLUSION AND FUTURE WORK

In this paper, we have revisited the BV metaphor for studying VA in tone mapped images in crowdsourcing experiment. BV based SM obtained with different parameters, based on perceptual features and

ecological experimental characteristics, have been compared to ET data. Results show that a bubble size of 2 visual degrees, with a blur level of 0.278° and a viewing duration per stimuli of 5 seconds lead to a good performance of the BV technique while the experimental conditions (content resolution, viewing angle and degrees per pixels) in ET and crowdsourcing experiments are comparable. However, some further works are needed to better understand the behavior of observers in BV experiments and the impact of content on this behavior in order to establish validated and efficient mouse-based metaphors for the crowdsourcing study of VA in different types of content for QoE applications.

6. ACKNOWLEDGMENT

The work in this paper was funded from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 765911, European Training Network on Real Vision project.

7. REFERENCES

- [1] Zoya Bylinskii, Tilke Judd, Aude Oliva, Antonio Torralba, and Frédéric Durand, "What do different evaluation metrics tell us about saliency models?," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 3, pp. 740–757, 2018.
- [2] Waqas Ellahi, Toinon Vigier, and Patrick Le Callet, "Can visual scanpath reveal personal image memorability? investigation of hmm tools for gaze patterns analysis," in *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2020, pp. 1–4.
- [3] James Laskey Nediya Daskalova Jeff Huang Alexandra Pappoutsaki, Patsorn Sangkloy and James Hays, "Webgazer: Scalable webcam eye tracking using user interactions," in *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*, 2016, p. 3839–3845.
- [4] Xucong Zhang, Michael Xuelin Huang, Yusuke Sugano, and Andreas Bulling, "Training person-specific gaze estimators from user interactions with multiple devices," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–12.
- [5] Anelise Newman, Barry McNamara, Camilo Fosco, Yun Bin Zhang, Pat Sukhum, Matthew Tancik, Nam Wook Kim, and Zoya Bylinskii, "Turkeys: A web-based toolbox for crowdsourcing attention data," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–13.
- [6] Peter O'Donovan, Aseem Agarwala, and Aaron Hertzmann, "Learning layouts for single-pagegraphic designs," *IEEE transactions on visualization and computer graphics*, vol. 20, no. 8, pp. 1200–1213, 2014.
- [7] Dmitry Rudoy, Dan B Goldman, Eli Shechtman, and Lihi Zelnik-Manor, "Crowdsourcing gaze data collection," *arXiv preprint arXiv:1204.3367*, 2012.
- [8] Nam Wook Kim, Zoya Bylinskii, Michelle A Borkin, Krzysztof Z Gajos, Aude Oliva, Fredo Durand, and Hanspeter Pfister, "Bubbleview: an interface for crowdsourcing image importance maps and tracking visual attention," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 24, no. 5, pp. 1–40, 2017.
- [9] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick, "Microsoft coco: Common objects in context," in *European conference on computer vision*. Springer, 2014, pp. 740–755.
- [10] Juan Xu, Ming Jiang, Shuo Wang, Mohan S Kankanhalli, and Qi Zhao, "Predicting human gaze beyond pixels," *Journal of vision*, vol. 14, no. 1, pp. 28–28, 2014.
- [11] Ming Jiang, Shengsheng Huang, Juanyong Duan, and Qi Zhao, "Salicon: Saliency in context," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1072–1080.
- [12] Waqas Ellahi, Toinon Vigier, and Patrick Le Callet, "Hmm-based framework to measure the visual fidelity of tone mapping operators," in *2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. IEEE, 2020, pp. 1–6.
- [13] Lukas Krasula, Manish Narwaria, Karel Fliegel, and Patrick Le Callet, "Influence of hdr reference on observers preference in tone-mapped images evaluation," in *2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE, 2015, pp. 1–6.
- [14] Tobii Pro Lab (Version 1.145) [Computer software], "Tobii pro ab," in *J. Danderyd, Sweden: Tobii Pro AB*, 2014.
- [15] Nam Wook Kim, Zoya Bylinskii, Michelle A Borkin, Krzysztof Z Gajos, Aude Oliva, Fredo Durand, and Hanspeter Pfister, "Bubbleview: an interface for crowdsourcing image importance maps and tracking visual attention," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 24, no. 5, pp. 1–40, 2017.
- [16] "Prolific," <https://www.prolific.co/>, Accessed: Oct 2020. [Online].
- [17] Kuan-Ta Chen, Chen-Chi Wu, Yu-Chun Chang, and Chin-Laung Lei, "A crowdsourcable qoe evaluation framework for multimedia content," in *Proceedings of the 17th ACM international conference on Multimedia*, 2009, pp. 491–500.
- [18] Manish Narwaria, Matthieu Perreira Da Silva, Patrick Le Callet, and Romuald PÉpion, "Effect of tone mapping operators on visual attention deployment," in *Applications of Digital Image Processing XXXV*. International Society for Optics and Photonics, 2012, vol. 8499, p. 84990G.