ULTRA-EYE: UHD AND HD IMAGES EYE TRACKING DATASET

Hiromi Nemoto, Philippe Hanhart, Pavel Korshunov, and Touradj Ebrahimi

Multimedia Signal Processing Group, EPFL, Lausanne, Switzerland

ABSTRACT

Due to the recent advances in ultra high definition (UHD) displays, UHD TV may replace HD TV in a near future. However, little is known about the effect of UHD content on human visual perception, specifically, on human visual attention, viewing strategies, and visual saliency. To help studying these properties of the human visual system and their dynamics when HD content is replaced with UHD content, a publicly accessible dataset is proposed, which is composed of 41 4K UHD and HD images with corresponding eye tracking information. The eye tracking information includes the fixation points and fixation density maps measured during extensive subjective experiments. In this paper, we describe the dataset in details, including the strategy for content selection, the eye-tracking experiments, and the computation of the fixation density maps.

Index Terms— Eye tracking, human fixation, visual attention, UHD images, dataset

1. INTRODUCTION

Understanding human attention patterns and viewing strategies for UHD content is important for developing efficient data compression algorithms and accurate objective quality metrics. Although a significant number of public image and video datasets for visual attention already exist [1], no dataset with eye tracking data is available for UHD content. However, without such data, it is hard to understand the impact of UHD on visual attention and whether it is significant for practical applications, e.g., computational models of visual saliency.

To answer this question, a publicly accessible dataset that can be used for research purposes is created[1], called Ultra-Eye, composed of 41 4K UHD and HD (downsampled from UHD) images with high variety of content (see examples in Figure 1).
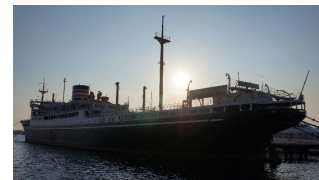
For each image, eye movement data is recorded via a set of subjective experiments using 20 naïve subjects and a professional eye tracker. From the eye movement data, the fixa-

Fig. 1: Examples of images from the Ultra-Eye dataset.

(a) *Presented picture*
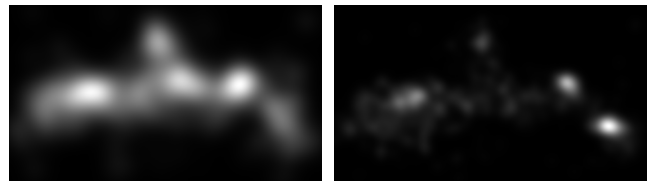
(b) *HD FDM*   (c) *UHD FDM*

Fig. 2: Examples of UHD and HD FDMs.

tion density maps (FDMs) for UHD and HD resolutions are computed. The resulted FDMs demonstrate differences between viewing strategies when watching UHD and HD contents, as illustrated by the examples in Figure 2.

In summary, the dataset provides the following:
1. UHD and HD images stored in TIFF format;
2. Lists of fixation points;
3. Fixation density maps.

2. CONTENT CREATION

The Ultra-Eye dataset consists of 41 still images (see examples in Figure 1) acquired using several cameras, including Sony DSC-RX100 II, Sony NEX-5N, FUJIFILM XF1, Olympus E-PL2, and RED SCARLET-X. Some of the images were obtained from Europeana internet portal[2]. All of the original images have resolution higher than 4K UHD, so they were

[1]The dataset can be downloaded here: http://mmspg.epfl.ch/ultra-eye

[2]Europeana: think culture, http://www.europeana.eu

cropped to 3840×2160 pixels of 4K resolution. An HD version of 1920×1080 pixels was also created for each image by using Lanczos downsampling filter.

To reflect the high descriptive nature of 4K UHD resolution, the images in the dataset contain small details with large variations, e.g., close ups with variable background, small objects in large landscapes, city skylines, etc. The images also cover a wide variety of scenes, including outdoor and indoor scenes, images of nature, people, animals, and historical scenes depicted in paintings. Since visual artifacts affect visual attention patterns [2], special care was taken to exclude from the dataset images impaired by blur, noise, optical flares, or other visible distortions.

3. EYE TRACKING EXPERIMENTS

The eye tracking experiments were conducted at the MMSPG test laboratory, which fulfills the recommendations for subjective evaluation of visual data issued by ITU-R [3]. The viewing conditions were set according to recommendation ITU-R BT.2022 [4]. Table 1 presents the detailed summary of the experiments.

Each subject participated to two sessions of 13 minutes each: one with UHD images and another with HD images, with a 15 minutes break in between. The order of these sessions was different for two halves of the 20 subjects, i.e., one half of the subjects watched the UHD images first followed by the HD images, while the order was reversed for the other half. This design was selected to reduce the influence of potential memory effects on visual attention from viewing the same contents twice. To reduce contextual effects, the stimuli orders of display were randomized applying different permutation for each subject.

4. FIXATION DENSITY MAPS

Fixation density maps (FDMs) are computed by convolving the recorded gaze points with a Gaussian filter, and then normalizing the result to values between 0 and 1. Only gaze points corresponding to fixation points are used for computing a FDM. Gaze points associated with saccades are not used in the computation. The eye tracking system used in our experiments (see Table 1 for details) automatically discriminates between saccades and fixations, based on the gaze velocity information. Blinks are also detected automatically by the eye tracking system based on the distance between the two eyelids of each eye. All detected saccades and blinks are excluded from the eye movement data and only the gaze points classified as fixation points are used further.

These points are then filtered with a Gaussian kernel to compensate the eye tracker inaccuracies and to simulate the foveal point spread function of the human eye. As suggested by Engelke *et al.* [5], the standard deviation of the Gaussian

Table 1: Overview of the eye tracking experiments.

Category	Details	Specification
Participants	Number ($\male$/$\female$)	20 (13/7)
	Age range (average age)	$18 - 27$ (23.8)
	Screening	Snellen and Ishihara charts
Viewing conditions	Environment	Laboratory
	Illumination	Low
	Color temperature	6500 [K]
	Viewing distance	1.12 [m]
	Task	Free-viewing
Display	Manufacturer	Sony
	Model	SRM-L560
	Type	LCD
	Size	56 [inch]
	Resolution	3840×2160 [pixels]
	Angular resolution	60 [pixel/degree]
Eye tracker	Manufacturer	Smart Eye
	Model	Smart Eye Pro 5.8
	Mounting position	0.7 [m] from the display
	Sampling frequency	60 [Hz]
	Accuracy	< 0.5 [degree]
	Calibration points	5 points on screen
Image presentation	Presentation order	Random
	Presentation time	15 [s]
	Grey-screen duration	2 [s]

filter used for computing the FDMs is set to 1 degree of visual angle, which corresponds to $\sigma = 60$ pixels in our experiments. This is based on the assumption that the fovea of the human eye covers approximately 2 degrees of visual angle.

5. CONCLUSION

This paper describes the Ultra-Eye public dataset, which consists of 41 images in UHD and HD resolutions with corresponding fixation points and fixation density maps obtained via eye tracking experiments. The dataset is intended for studying human visual attention and viewing strategies related to UHD and HD content.

6. REFERENCES

[1] S. Winkler and R. Subramanian, "Overview of Eye tracking Datasets," in *Fifth International Workshop on Quality of Multimedia Experience (QoMEX)*, July 2013, pp. 212–217.

[2] J. A. Redi, H. Liu, R. Zunino, and I. Heynderickx, "Interactions of visual attention and quality perception," in *IS&T/SPIE Electronic Imaging 2011 and Human Vision and Electronic Imaging XVI*, Jan. 2011, vol. 7865.

[3] ITU-R BT.500-13, "Methodology for the subjective assessment of the quality of television pictures," ITU, Jan. 2012.

[4] ITU-R BT.2022, "General viewing conditions for subjective assessment of quality of SDTV and HDTV television pictures on flat panel displays," ITU, Aug. 2012.

[5] U. Engelke, A. Maeder, and H. Zepernick, "Visual attention modelling for subjective image quality databases," in *IEEE International Workshop on Multimedia Signal Processing (MMSP)*, Oct. 2009, pp. 1–6.