https://doi.org/10.2352/ISSN.2470-1173.2017.14.HVEI-148

Developmental Changes in Ambient and Focal Visual Processing Strategies

Onkar Krishna[1], Toshihiko Yamasaki[1], Andrea Helo[2], Rämä Pia[2], Kiyoharu Aizawa[1]; The University of Tokyo, Tokyo, Japan[1]; Paris Descartes University, Paris, France[2]

Abstract

The relationship between fixations and saccades that characterize eye movements have suggested the existence of two scene processing strategies- ambient and focal - the former is believed to capture the context of the scene whereas the latter helps in the detailed examination of the scene. Until recently, detailed analysis of these processing strategies have been performed only in young adults. Recently, however, these modes were found to exist in the viewing behavior of children as well. In the current study, we investigated two undiscovered aspects related to these processing modes- first, we studied the combined impact of visual processing modes and observer's age on the distribution of gaze locations, and second, we investigated how bottom up features of a scene influence the gaze behavior during ambient and focal modes across age groups. These analysis were performed over the eye tracking data collected over 50 observers of different age groups while viewing naturalistic scenes. Exploratiweness and bit rate measures were developed to investigate the changes in gaze distribution. The result showed that an observer is more explorative during the focal mode than during the ambient processing mode, while the information processing rate follows a reverse trend. The influence of bottom up features was investigated using the area under the curve (AUC) metric and it was found that the bottom-up influence was more dominant during focal than the ambient mode for all age groups. We also investigated whether human face attracted attention and gaze differently during ambient and focal mode. The result showed that face were equally gazed during both processing modes.

Introduction

Human's gaze behavior during scene viewing is mainly characterized by fixation duration and saccade amplitude at fixated locations. The systematic tendencies of gaze deployment were initially identified in studies of Buswell (1935) [20], and Antes (1974) [13], where it was observed that fixation durations increases and saccade amplitudes decrease over viewing time. More recently Unema et al., (2005) [8], reported that shorter fixations were associated with longer saccades and conversely, longer fixation was associated with shorter saccades, suggesting the existence of two processing modes: ambient and focal modes, respectively. Ambient mode generally answers to the question of 'where' in the scene and the focal mode is concerned with 'what' is the nature of objects being viewed in the scene [2].

Developmental aspects of the eye movement behavior have been extensively studied in past ([3], [4], and [5]). However, development of scene perception is relatively unexplored but it has been recently addressed by Helo et al., 2014, [1]. Their results suggested that as the age of an observer increases, fixation duration decreases and saccade amplitude increases during scene exploration. In addition, the study revealed the existence of ambient and focal processing modes in as early as 2 years of age. In another study, Alper et al., 2010, [17] found that the bottom-up features of the images guide the fixation landings more in younger than children. Even though these studies investigated developmental changes in scene viewing behavior in terms of considering the fixation duration and saccade amplitudes, they did not studied the distribution of gaze locations. Moreover, the parametric measures used in these studies do not suffice to make comparative comments on the nature of explored region across different age groups.

The present study aimed to investigate age-associated inter-individual differences in the gaze behavior during ambient and focal visual processing modes. The three major objectives of the current study were, first, to investigate the maturation of gaze distribution during ambient and focal modes. Exploratiweness and bit rate analysis metrics were used to analyze the gaze distribution. Second, we investigated the influence of bottom-up features of the scene during ambient and focal processing in different age groups. Third, we studied weather a face or a head (manually labeled) in a scene attracts gaze differently during scene viewing across age groups and two modes of viewing.

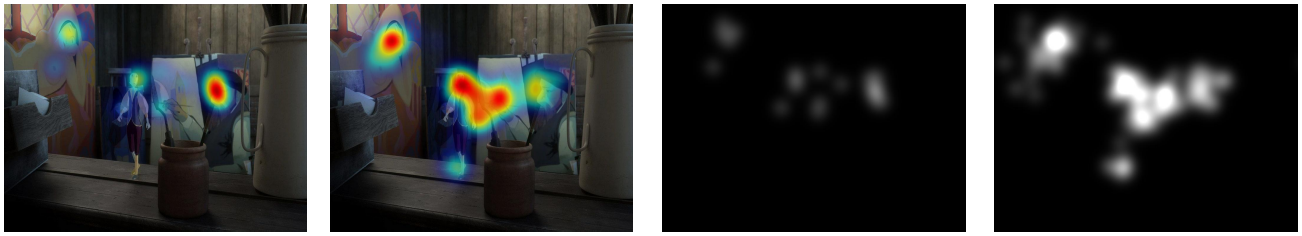
Method and Materials

Eye tracking data

Subjects and stimuli:- We analyzed the eye-tracking dataset collected in [1].We used eye-tracking data of 50 observers from different age groups. All the observers had normal vision or corrected to normal vision. Based on the participant's age, they were assigned in four age groups 4-6 year, 6-8 year, 8-10 year and adult (mean age 29 years). We use 4 year, 6 year, 8 year and young adult to refer to these groups in order. The experiment was conducted on images of the dimension of 1024×764 pixels. These images were taken from children's books and movies and characterized to have an eventful background.

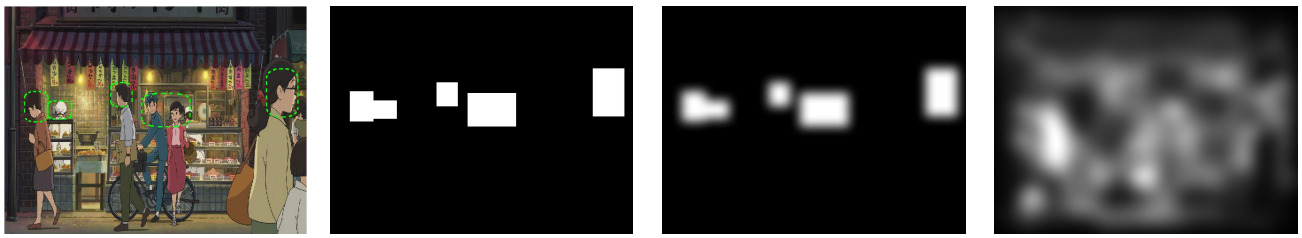
Apparatus and Procedure:- Remote eye tracking system Eyelink 1000 with a sampling rate of 500Hz was used to measure eye gaze and provided the raw data which was sampled to obtain fixations and saccades. Pictures were viewed from distance of 60 cm on the screen having a resolution of 1024×728 pixels. Each stimulus was presented for 10s in each trail.

Data Representation:- For a given age group we used classified fixation points by all the observers to generate the saliency map and heat map for all images as shown in Figure 1. Ambient and focal human saliency maps were obtained by convolving a Gaussian filter across the fixation locations, as in [13].



(a) Ambient heat map *(b) Focal heat map* *(c) Ambient saliency map* *(d) Focal saliency map*

Fig. 1: Data Representation:- *Ambient and focal saliency maps obtained by convolving the Guassian at fixation locations*



(a) Labelled head locations *(b) Binarized map* *(c) Head map* *(d) Bottom-up Saliency map*

Fig. 2: Head map and Saliency map:- *Head map generated by Gaussian convolution of the binarized map obtained from the manually labelling of the head locations on the data. Bottom-up features are represented by bottom-up saliency map.*

Data Analysis

In this section, we quantitatively analyze ambient-focal dichotomy for observers across age groups and its impact on gaze locations in terms of explorativeness and bit rate analysis. Consecutively, we studied the influence of bottom-up features during two scene exploration strategies and finally investigated the role of human heads in attracting observer's attention during the two modes.

Quantitative measures to analyze developmental changes in gaze distribution during ambient-focal processing modes

Here we introduce the parameters used for the comparative analysis of ambient and focal visual processing modes and subsequently investigate the developmental changes in two viewing strategies for scene viewing. We introduce two metrics to understand the role of the ambient and focal mode in fixation landings across age groups: (a) Explorativeness, which quantizes the distribution of explored region during ambient and focal visual processing modes of observers from different age groups. (b) bit rate metric to quantize the rate at which observers process the scene i.e. the information processing rate during ambient and focal processing modes.

(a) Explorativeness

Explorative behavior of an observer is his tendency to probe a given scene. In general observers who tend to focus only on some locations in the scene are less explorative than those who try to explore the entire scene. The observer's probing tendency differs drastically during ambient and focal visual processing strategies. This difference becomes even more prominent when we observer the gaze behavior of observers from different age groups.

Ambient and focal saliency maps are plausible map to be used to analyze explorative behavior of an observer as the human saliency maps for the two modes differ with age. First order entropy of ambient and focal saliency maps are used to quantify the explorativeness of ambient and focal modes across age groups. Formulation of the explorativeness is as following-

$$H(A_i^g) = \sum_l h_{A_i^g}(l) * log(L / h_{A_i^g}(l)) \quad (1)$$

$$H(F_i^g) = \sum_l h_{F_i^g}(l) * log(L / h_{F_i^g}(l)) \quad (2)$$

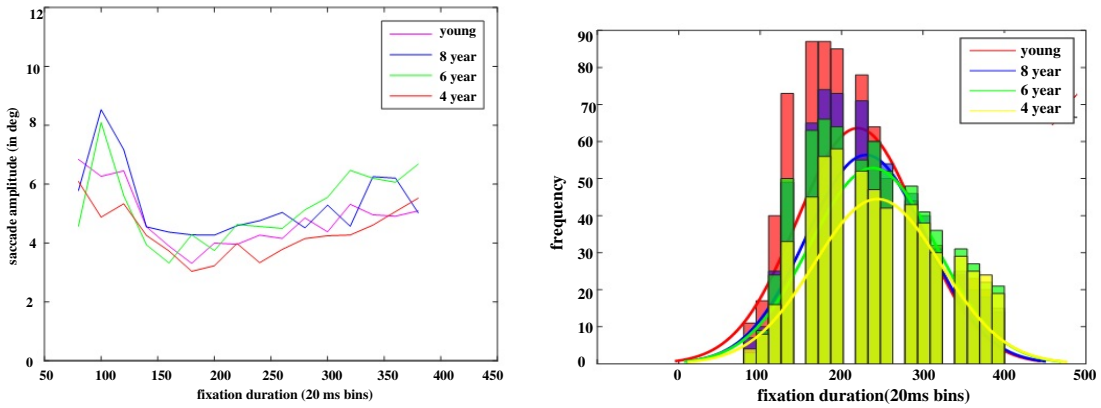
Where A_i^g and F_i^g are the ambient and focal saliency maps respectively, generated for i^{th} image over fixation landings of all observers in a group g. $h_{A_i^g}(l)$ and $h_{F_i^g}(l)$ are the histogram entries of intensity value l for saliency maps A_i^g and F_i^g respectively, L is the total number of pixels in A_i^g. The more scattered a saliency map is i.e., the more the number of non zero pixels in the map, the more is its entropy.

In the context of viewing behavior, higher entropy corresponds to the higher explorative viewing behavior of observer as they have more number of scattered saliency points in the scene. Similarly, lower entropy corresponds to the less explorative behavior.

(b) Bit rate analysis

Exporativeness gives an idea about explored regions but does not consider the time taken to explore the corresponding regions during ambient and focal modes. The study reported in [8] reveals that ambient fixations are associated with significantly shorter fixation duration than the focal fixations which suggests that the visual information processed during ambient and focal modes accord to different time intervals.

The fundamental difference in ambient-focal strategies suggests the requirement of a metric which can measure the rate of visual information processing during ambient and focal scene processing. The bit/second metric is developed for this purpose. It is defined as the ratio of a total number of bits required to represent the ambient and focal saliency maps of an image to the total time associated with ambient and focal fixations.



(a) *(b)*

Fig. 3: (a):Saccade amplitude as function of fixation duration:- *Irrespective of the age groups shorter fixations are associated with the longer saccade and relatively longer fixations associated with shorter saccades* ***(b): Frequency plot of the number of fixation in each bin:-*** *young observer's distribution is more concentrated towards the average fixation of the group than children which indicates that ambient and focal tendency changes with age*

Influence of bottom-up features on the ambient and focal processing modes

Bottom-up features of a scene include features such as color, intensity, and orientation. These are best represented by the Ittis saliency map [19], which extracts these features in parallel by a set of the linear center surrounded operations similar to the visual receptive field, and then combines their normalized value at the later stage to obtain the conspicuous location of an image.

To study the influence of the bottom up features over the two visual processing modes, the bottom-up saliency maps are first thresholded to T different levels representing the different percentage of the most salient area of the map. These are then used to predict ambient and fixation landings for different age groups and the prediction performance is measured using AUC score. [11].

Role of human heads in seeking observer's attention during the two modes

A study in [10] on eye movement data of fifteen observers (18-35 years) over 1003 natural images revealed that observers frequently fixated on faces in scenes. This result helps in the development of a better computational model of visual attention by considering faces in the generation of feature maps. In order to develop age adopted saliency models, it gets important to investigate the role of human heads/faces across age groups. The final objective of this study is to investigate how the human's head (manually labeled) attracts observer's attention during two processing strategies. For the same, we generated a head map for the scene. Further, these maps are used to evaluate the prediction performance for ambient and focal fixation landings.

To generate head maps first the head locations in a scene are manually labeled for all the images in the dataset, further these manually labeled locations are binarized, and finally Gaussian blurring is used to minimize the effect of the sharp boundaries around the head locations in the binary map. The results of these steps for a sample image are shown in Figure 2. AUC based measures are used for this purpose.

Data preprocessing: fixation classification: -

In order to understand the maturation of fixation distribution during two modes, fixations were classified in ambient and focal following the same steps as reported in [1][8]. The focal and ambient population of the fixations is decided based on the relationship between fixation duration and saccade amplitude. To examine this relationship we plotted average value of saccade amplitude within bins of the size 20 ms for all the fixations up to 400 ms. It can be clearly seen from Figure 3(a) the fixations across all age groups can be broken down into two distinguishable subsets, one characterized by significantly larger saccades compared to the other. Further to statistically validate the results, MANOVA analysis was performed. The results reported this effect significant across all the age groups (**adults-** $F(1,14) = 15.31, p < 0.001$, **8 year-** $F(1,14) = 7.46, p = 0.016$, **6 year-** $F(1,14) = 5.51, (p < 0.043)$, **4 year-** $F(1,14) = 6.07, p < 0.027$).

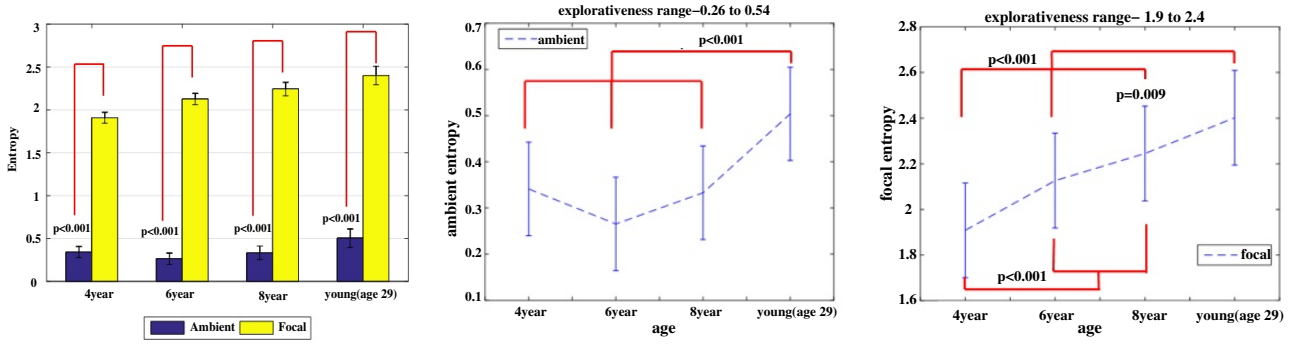
A bar graph of a number of fixations in each bin was plotted to understand the distribution of a number of fixations across bins. As can be seen in Figure 3(b), the fixations distribution gradually become less spread with increasing age which implies, the fixations converge towards the average value of fixation duration with increasing age of an observer. In other words, the percentage of fixations with exceptionally high and low fixation duration decreases with increasing age.

Results

We identified fixations in two classes across age groups and formalized the metric to quantify the gaze distributions in terms of explorativeness and bit rate. Here we report the results of these analyses in order to understand developmental changes occurring in these two processing modes.

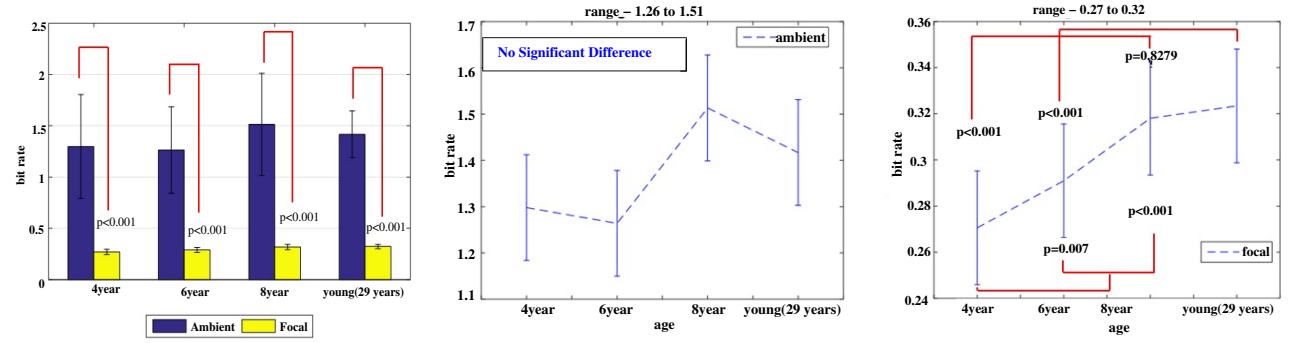
Ambient and Focal Explorativeness

Explorativeness was measured for different age groups during different processing modes. It was found that the explorative behavior does not only vary for focal and ambient processing modes but it is also influenced by an observer's age. The comparative results of explorativeness are the following:



(a) *(b)* *(c)*

Fig. 4: *Ambient and focal exploratiyeness is measure by the entropy of the respective saliency maps.* ***(a) Ambient vs. focal:*** *The bar graph shows the focal exploratiyeness is significantly higher than ambient.* ***(b) Ambient exploratiyeness vs age:*** *Age impact is not significant for ambient exploratiyeness during childhood.* ***(c)Focal exploratiyeness vs age:*** *fixations distribution are significantly different among age groups.*



(a) *(b)* *(c)*

Fig. 5: *Results of the bit rate analysis* ***(a) Ambient vs. focal bit rate:*** *Ambient information processing rate is significantly higher than the focal.* ***(b) Ambient rate vs age:*** *There is no significant impact of the age during ambient scene processing rate.* ***(c) Focal rate vs. age:*** *The focal rate of scene processing significantly increases with observer's age.*

Focal is more explorative than Ambient:- Exploratiyeness measure showed that observers were more explorative during the focal mode than in the ambient mode. The same observation was found in all age groups. Paired t-test showed that the exploratiyeness measure was statistically significant between the ambient and focal modes in all age groups ($p < 0.001$-for all age groups) as in Figure 4(a).

Focal exploratiyeness increases with age:- The result showed that focal exploratiyeness increased with age. Exploratiyeness of adults were significantly higher than the children ($p < 0.001$) (ANOVA analysis results reported in Figure 4(b)). However the changes in explorative behavior for ambient processing mode was not significant during childhood as can be seen in Figure 4(c).

Ambient and Focal Bit Rate Analysis

Having analyzed the explorative behavior of observers, their capability of quickly grabbing the scene context during ambient mode and detailed processing of scene during focal mode was measured in terms of Bit rate metrics. Similarly to the exploratiyeness analysis comparisons was performed between two visual processing modes and over different age groups.

Information processing rate is higher in ambient mode:- The rate of the visual information processing was higher in ambient mode than in the focal mode (ambient range- 1.26 to 1.51, focal range- 0.27 to 0.32) and it was true for all age groups. The statistical values are indicated in Figure 5(a).

Focal rate increases with observer's age:- It is also observed that the rate of visual information processing during ambient mode was same across age groups. On the contrary, the information processing during focal processing increased with age. The statistical values are indicated in Figure 5(b), and (c).

Role of the bottom-up features and head locations in attracting the gaze during ambient and focal mode

Area under the curve (AUC) metrics was used to measure the bottom-up influence during ambient and focal scene processing modes. Bottom-up saliency map was used to report the prediction performance of the ambient and focal fixation landings. Our experiment indicated that the focal mode was more bottom-up than ambient for all age groups, i.e., bottom-up saliency maps are more suitable to predict the focal maps as compared to the ambient maps. In order to validate these results the paired t-test was performed between AUC values of the ambient and focal processing for all the age group; statistical values for which are indicated in Figure 5.

Head maps were used to investigate the influence of head locations during ambient and focal modes among different age

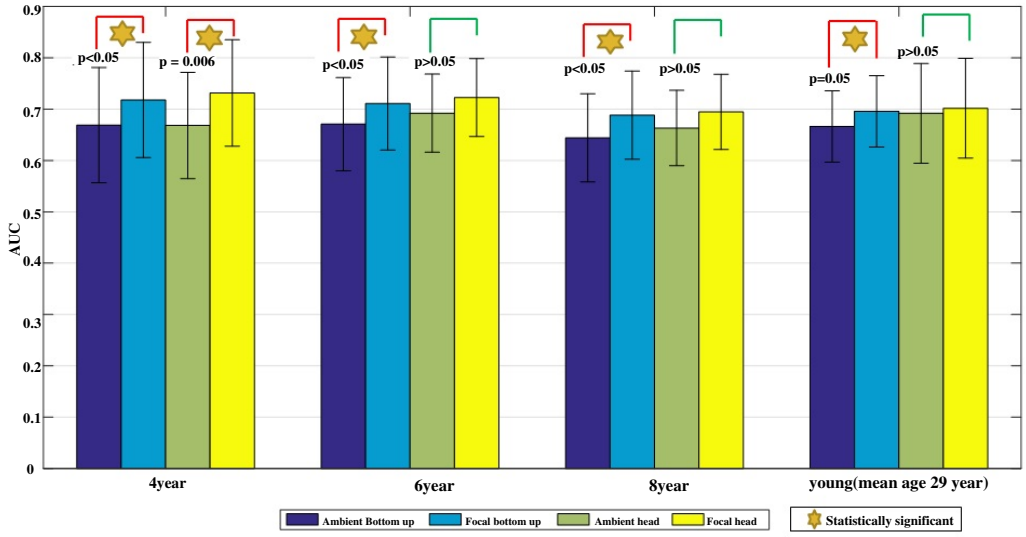


Fig. 6: *Bottom-up and head map influence over ambient and focal modes for different age. AUC score 0.5 is for random performance*

groups by measuring the AUC values. The result revealed that the head locations in scenes attracted the observer's attention equally under the two processing modes for all age groups, except for the observers of four years of age. The statistical significance values are indicated in Figure 6.

Discussion and Conclusion

In the current work, our aim was to investigate developmental changes in gaze distribution during the two visual processing modes, ambient and focal, across four age groups. The explorativeness and bit rate analysis metrics were formulated to quantify the gaze behavior of observers. Explorativeness measures the spread of fixations while bit rate helps to quantify the fast grabbing and detailed analysis capability of observers during ambient and focal modes, respectively.

Our results revealed that an observer is more explorative during the focal than the ambient mode of scene viewing and that the explorativeness during the focal mode increases with age of an observer. Similar results were reported in [19] where development of visual cognitive functions during scene viewing was analyzed by examining the exploratory eye movements of 4 to 16 years old observers using two metrics - a total number of gaze points (TNGP) and eye-scanning length (TESL) of gaze points. These result suggested that both TNGP and TESL increases with age suggesting the useful biologic markers for estimating the development of visual cognitive function

Ambient maps are characterized by fewer fixations with shorter fixation duration and the converse is true for focal maps([1], [2], and [8]). However, there is no evidence to the relationship between the proportion of fixations and total fixation duration in two modes. Bit-rate metric was developed to parametrize this relation and to examine the developmental changes in this behavior. The results of our study suggested that an observer has higher information processing rate during the ambient mode as compared to the focal mode and this is true for all age groups. Further, information processing rate in focal mode increases with age.

Next major contribution of our work was to demonstrate the impact of bottom-up features over eye movements during ambient and focal modes across the age groups. The results suggested that focal mode is more bottom-up than ambient across all age groups. The results of this study are in agreement with previous findings reported in [2], which also suggested that focal visual processing mode is more bottom-up than ambient for young adults. Similarly, the influence of bottom-up features was investigated by [1] during the early and late phases of viewing, which showed that early phase was more influenced by bottom-up than late phase in older groups.

Finally, we investigated whether an observer is more likely to allocate attention to face or head locations during the ambient or focal processing mode. The results suggested that except for the youngest children (4-year) head locations had a similar impact on attracting the gaze locations during ambient and focal modes of viewing. One similar study reported in [7] revealed that faces attract gaze independent of a task.

References

[1] A. Helo, S. Pannasch, L. Sirri, P. Rama, The maturation of eye movement behavior: Scene viewing characteristics in children and adults. Vision Research, 103, 2014, pp. 83-91.

[2] B. Follet, O. Le Meur, T. Baccino, New insights into ambient and focal visual fixations using an automatic classification algorithm. I-Perception, 2, 592610.

[3] O. Braddick, J. Atkinson, Development of human visual function, Vision Research, 51 (13) (2011), pp. 15881609.

[4] A. Roucoux, C. Culee, M. Roucoux, Development of fixation and pursuit eye movements in human infants, Behavioural Brain Research, 10 (1) (1983), pp. 133139

[5] A. Chandna, Natural history of the development of visual acuity in infants, Eye, 5 (1991), pp. 2026.

[6] Kosslyn SM, Brown HD, Dror IE, Aging and scope of visual attention, Gerontology 45, 102109, (1999).

[7] M. Cerf, E. P. Frady, C. Koch, Faces and text attract gaze independent of the task: Experimental data and computer model, J. Vis., vol. 9, no.

12, pp. 10, 2009.
[8] P.J.A. Unema, S. Pannasch, M. Joos, B. M. Velichkovsky, Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. Visual Cognition, 12, 473494, 2005.
[9] Emergent Techniques for Assessment of Visual Performance, National Research Council (US) Committee on Vision. Washington (DC): National Academies Press (US); 1985.
[10] Tilke Judd, Krista Ehinger, Frédo Durand, and Antonio Torralba, Learning to Predict Where Humans Look, IEEE International Conference on Computer Vision (ICCV), 2009
[11] O. Le Meur and T. Baccino, Methods for comparing scanpaths and saliency maps: strengths and weaknesses, Behavioral Research, 2012.
[12] T.Judd, K.Ehinger, F.Durand, A.Torralba, Learning to predict where humans look, Proceeding International Conference Computer Vision, pp.1-7, 2009.
[13] J R. Antes, The time course of picture viewing, Journal of Experimental Psychology. 1974;103:6270.
[14] S. Pannasch, J. R. Helmert, K. A. K. Roth, H. Walter, Visual Fixation Durations and Saccade Amplitudes: Shifting Relationship in a Variety of Conditions, Journal of Eye Movement Research, 2008; 2:119.
[15] C B. Trevarthen, Two mechanisms of vision in primates, Psychologische Forschung. 1968; 31:299337. doi: 10.1007/BF00422717.
[16] M. Mills, A. Hollingworth, M.D. Dodd, Examining the influence of task set on eye movements and fixations Journal of Vision, 11 (8) (2011), pp. 115.
[17] A. Acik, A. Sarwary, R. Schultze-Kraft, S. Onat, P. Knig, Developmental changes in natural viewing behavior: Bottom-up and top-down differences between children, young adults and older adults. Frontier in Psychology,1:207,(2010).
[18] C. Egami, K. Morita, T. Ohya, Y. Ishii, Y. Yamashita, T. Matsuishi Developmental characteristics of visual cognitive function during childhood according to exploratory eye movements Brain and Development, 31(10)(2009), pp. 750757.
[19] L. Itti, C. Koch, and E. Niebur A model of saliencybased visual attention for rapid scene analysis, IEEE Patt. Anal. Mach. Intell., 20(11): 1254-1259, November 1998.
[20] G. Buswell, How people look at picture. Chicago: University of Chicago Press, 1935.

Author Biography

Onkar Krishna *received his B.Tech. and M.Tech. in Computer Science from IILM, Noida, India in 2009 and IIIT, Jabalpur, India in 2012 respectively. He is currently pursuing his Ph.D. from the Department of Information and Communication Engineering, The University of Tokyo, Japan from 2015. His research interests include Computer Vision, Image Processing, Signal Processing and Audio watermarking and is currently engaged in personalization of saliency model.*

Prof. Kiyoharu AIZAWA *received the B.E., the M.E. and the Dr.E. in electrical engineering all from the University of Tokyo in 1983, 1985, 1988 respectively. He is currently a Professor at the Department of Information and Communication Engineering of the University of Tokyo. His main areas of research interest are image coding, processing and multimedia applications, and he is currently engaged in multimedia lifelog and other multimedia related subjects. Dr. Aizawa is currently an Senior Associate Editor of IEEE Trans. Image Processing and is on Editorial Board of ACM TOMM.*

Dr. Toshihiko YAMASAKI *received the B.S. degree in electronic engineering, the M.S. degree in information and communication engineering, and the Ph.D. degree from The University of Tokyo in 1999, 2001, and 2004, respectively. He is currently an Associate Professor at Department of Information and Communication Engineering, The University of Tokyo. His research interests include 3D processing, image processing, multimedia, computer vision, pattern recognition and VLSI design and is currently engaged in Big Multimedia Data Analysis and User Navigation, Object/action Recognition and Machine Learning, Optimization and High-Speed Processing and 3D Image/Video Processing. He is a member of IEEE, ACM, ITE, IEICE, IPSJ, and APSIPA.*

Andrea Helo *is a Ph.D. student at the Laboratoire Psychologie de la Perception, Universit Paris Descartes , Paris since October 2012. Her research interests include Language and vision Development: Ocular movement control during scene perception and semantic organization in Infants.*

Dr. Pia Rämä *received her M.A. and Ph.D. in pyschology from Univerity of Helinsky in 1994 and 1998 respectively. She was a post doctoral fellow at Institute of Biomedicine, University of Helsinki and Department of Psychological and Brain, Sciences, Johns Hopkin University from 1998-99 and 1999-2001 respectively. She was a research fellow from 2001-06 at Cognitive Brain Research Unit, University of Helsinki and from 2007-09 at Department of Pyschology at University of Nice Sophia Antipolis, France. She is currently a Research Scientist at Laboratoire Psychologie de la Perception , Paris Descartes University, Paris. Her research interest include language development, lexical-semantic processing, bilingualism, scene perception, attention, working memory, prefrontal cortex.*