Billboard Saliency Detection in Street Videos for Adults and Elderly

Onkar Krishna[1, 2]; Kiyoharu Aizawa[1]
[1]The University of Tokyo, Japan, [2]NTT Communication Science Laboratories, Japan

Research Objective

An age-adapted video saliency algorithm that can predict the relative saliencies of billboards in a street video for adult and elderly

Contributions

1. An eye-gaze data collected for adults and elderly while viewing a street video in free-viewing and task viewing.
2. Suitable metrics to analyze the collected data to reveal the age-related differences.
3. An age-adapted video saliency algorithm which predicts the relative saliencies of the billboards

Materials

Participants and Stimuli:
- 30 participants - 15 participants - mean age 24 and 73.1 each
- Stimuli - 2 min 30 sec street video full of restaurants.
- Tobii x2-60 eye tracker used for recording the eye-gaze data,
- fixations and saccades were detected by Tobii fixation filter.

Procedure and Task:
- Participants were divided into two groups of 15 (mix of both age groups)
- Group 1 watched the video in free viewing, and group 2 watched it in task viewing mode
- Task – search a place for lunch

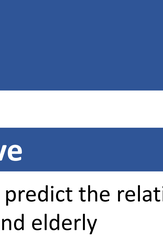

Figure 1: Illustration of the experiment setup

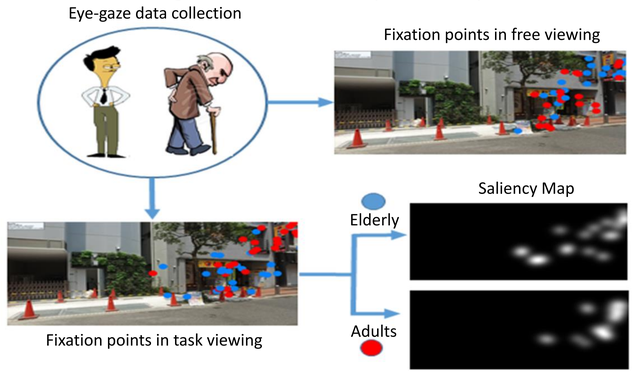

Figure 2: Process of generating fixation map and saliency map

Methodology

Figure 3: Visualization of different tendencies of gaze landings around the billboards in free viewing for adults (first row) and elderly (second row) participants

Figure 4: Few instances of manually labelled billboards (only restaurants).

Metrics used to quantify the age-related differences
1. Total fixation count
2. Explorativeness
3. Center bias.

Explorativeness :

$$H(I_{A/E}^{task}) = \sum_{l} h_{I_{A/E}^{task}}(l) * log(L \,/\, h_{I_{A/E}^{task}}(l)) \qquad (1)$$

$$H(I_{A/E}^{free}) = \sum_{l} h_{I_{A/E}^{free}}(l) * log(L \,/\, h_{I_{A/E}^{free}}(l)) \qquad (2)$$

- Result :
  - Free viewing - adults are significantly more explorative than elderly
  - Task viewing - both age groups showed similar tendency.

Center bias - elderly have higher center bias than adults in both modes

Figure 5: Annotated billboards in the street video

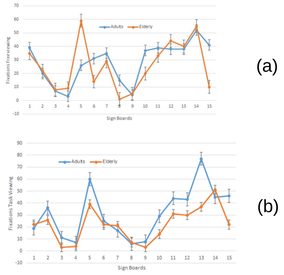

Figure 6: Billboards ranking in (a) free viewing and (b) task viewing

Saliency Model

We have developed a learning based model, based on idea proposed in a previous study [6] with following changes to adapt with the age-related differences:
- Selected optimum set of scales for different features (bottom, mid, top-level features) extracted for adults and elderly.
- The models weights for feature combination are learned for the adults and elderly participant's gaze separately.
- Center-maps are tuned according to age-related differences in center-bias tendency.

Experiment setup:
- The top 10% strong positive from top 5% of the ground truth saliency map and similarly 10% strong negative pixels from bottom 20% of the map were selected
- For a given set of features and levels (positive and negative samples) SVM is used to learn the optimal weights i.e., model parameters for adults and elderly.

$$SM_A(I_i) = w_A F^T(I_i) + b_A \qquad (3)$$

$$SM_E(I_i) = w_E F^T(I_i) + b_A \qquad (4)$$

Table 1: The prediction accuracy (AUC score) billboards for adults and elderly participants

		Free Viewing				Task Viewing			
Board Saliency		Highly Salient		Least Salient		Highly Salient		Least Salient	
Board number (adults)		1	14	4	9	5	13	8	9
Board number (elderly)		5	14	8	9	5	14	4	9
Itti [1]	adults	0.87	0.65	0.78	0.60	0.76	0.69	0.72	0.72
	eldery	0.75	0.67	0.73	0.63	0.75	0.77	0.60	0.68
GBVS [2]	adults	0.83	0.71	0.83	0.57	0.78	0.76	0.63	0.58
	eldery	0.79	0.70	0.58	0.57	0.71	0.72	0.81	0.65
s_map [3]	adults	0.74	0.73	0.82	0.58	0.70	0.73	0.60	0.57
	eldery	0.72	0.71	0.62	0.61	0.83	0.71	0.70	0.71
m_map [4]	adults	0.67	0.64	0.65	0.68	0.78	0.73	0.76	0.59
	eldery	0.65	0.60	0.69	0.56	0.69	0.70	0.70	0.59
e_map [5]	adults	0.71	0.65	0.71	0.55	0.65	0.66	0.56	0.53
	eldery	0.66	0.66	0.54	0.58	0.58	0.78	0.58	0.61
Ours	adults	0.83	0.79	0.67	0.65	0.80	0.76	0.69	0.63
	eldery	0.78	0.72	0.63	0.68	0.75	0.77	0.72	0.69

Conclusions

- Adults and elderly while viewing street videos have difference gaze landings during free viewing and task viewing as has been verified by our metrics.
- The proposed model outperforms others in predicting saliency for billboards in paved areas of street videos for different age groups.

Acknowledgement

We are thankful to VTEC laboratories Inc. for providing the video used in this study.

Contact

Onkar Krishna
Media Recognition Research Group,
NTT Communication Science Laboratories,
NTT Corporation, Japan
onkar@lab.ntt.co.jp

References

1. Laurent Itti, Christof Koch, and Ernst Niebur. 1998. A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on pattern analysis and machine intelligence 20, 11 (1998), 12541259.
2. Jonathan Harel, Christof Koch, and Pietro Perona. 2007. Graph-based visual saliency. In Advances in neural information processing systems. 545552.
3. Yuming Fang, Zhenzhong Chen, Weisi Lin, Chia-Wen Lin: Saliency Detection in the Compressed Domain for Adaptive Image Retargeting. IEEE Transactions on Image Processing 21(9): 3888-3901 (2012)
4. Yuming Fang, Weisi Lin, Zhenzhong Chen, Chia-Ming Tsai, and Chia-Wen Lin, A Video Saliency Detection Model in Compressed Domain. IEEE Trans. Circuits Syst. Video Techn. 24(1): 27-38, 2014.
5. Fang, Y., Wang, Z., Lin, W., and Fang, Z. (2014). Video saliency incorporating spatiotemporal cues and uncertainty weighting. IEEE Transactions on Image Processing, 23(9), 3910-3921.
6. Tilke Judd, Krista Ehinger, Frdo Durand, and Antonio Torralba. 2009. Learning to predict where humans look. In Computer Vision, 2009 IEEE 12th international conference on. IEEE, 21062113.
7. Krishna, O., Aizawa, K., and Reimerth, S. (2018). "Signboard Saliency Detection in Street Videos." Acoustics, Speech and Signal Processing (ICASSP), IEEE International Conference on. IEEE, 2018.
8. Krishna, O., Helo, A., Rämä, P. and Aizawa, K., 2018. Gaze distribution analysis and saliency prediction across age groups. *PloS one*, *13*(2), p.e0193149.