Unraveling Human Perception of Facial Aging using Eye Gaze

Daksha Yadav[1], Naman Kohli[1], Ekampreet Kalsi[2], Mayank Vatsa[3], Richa Singh[3], Afzel Noore[4]
[1]West Virginia University, [2]University of Florida, [3]IIIT-Delhi, [4]Texas A&M University-Kingsville
[1]{dayadav, nakohli}@mix.wvu.edu, [2]ekam2148@ufl.edu, [3]{mayank, rsingh}@iiitd.ac.in,
[4]afzel.noore@tamuk.edu

Abstract

Continuous efforts are being made to understand human perception network with the purpose of developing enhanced computational models for vision-based tasks. In this paper, we utilize eye gaze as a medium to unravel the cues utilized by humans for the perception of facial aging. Specifically, we explore the tasks of face age estimation and age-separate face verification and analyze the eye gaze patterns of participants to understand the strategy followed by human participants. To facilitate this, eye gaze data from 50 participants is acquired using two different eye gaze trackers: Eye Tribe and GazePoint GP3. Comprehensive analysis of various eye movement metrics is performed with respect to different face parts to illustrate their relevance for age estimation and age-separated face verification tasks.

1. Introduction

Human facial aging is a complex natural process. During the lifetime of an individual, noticeable variations in the facial appearance can be observed and several factors such as bone growth, gender, and ethnicity [1] contribute to how an individual ages, making it a very individualistic function. As illustrated in Figure 1, researchers [2, 3] have observed that during the formative years of an individual, transformations in the facial shape are prominent while in later years, textural variations such as wrinkles and pigmentation are more noticeable. These inherent variations allow humans to decode the age of an individual by merely looking at their face.

Correctly verifying age-separated faces is required in several critical scenarios such as law enforcement, airport security, and border control. In these application domains, officers match the images in the presented ID or passport to the person in front of them. Several psychological and perception based studies have analyzed the performance of human participants for the task of age estimation [4, 5] and face verification [6, 7]. It has been established that humans can perform this task accurately even in adverse scenarios

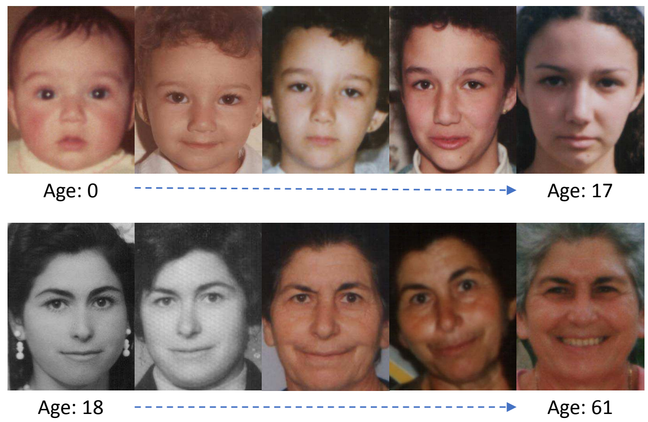


Figure 1: Illustrating variations in facial variations due to aging. The top row demonstrates changes in the facial appearance of an individual from age 0 to 17. The bottom row illustrates facial feature variations of another individual from age 18 to 61. Image source: [10]

such as occluded or degraded faces. Unlike facial age estimation, the evaluation of human expertise in matching age-separated face images is a relatively lesser explored area. Even though automatic face verification algorithms struggle with this covariate [8, 9], humans are fairly accurate in determining if two age-separated face images belong to the same individual or not [4]. Yadav et al. [4] conducted the first study on recognizing age-separated face images by participants.

All the above-mentioned studies utilize participants' responses to categorical questions to learn about the human expertise in these tasks. An instance of such a question is: *Determine if the two images belong to the same individual: Yes or No*. However, these responses may not provide detailed insight of the innate network present in human beings which is responsible for this skill. Contrastively, studies have demonstrated that what an individual is looking at can provide insight into the cognitive processes. Thus, eye gaze tracking can reveal information about the individual's attention for a given visual stimuli [11]. Computation of different eye gaze analysis based metrics such as fixations may provide additional cues about the cognitive load. Eye tracking technology has been successfully deployed to learn an individual's intent for various application such as user-specific advertising [12] and developing user-friendly interfaces [13, 14]. With this motivation, we utilize eye gaze tracking to unravel how human perceive facial aging and investigate different face parts to understand their significance for the perception task.

Visual cognition based studies have established that attentive mechanisms in human brain supervise the gaze only on the salient parts of the stimuli [15]. Therefore, in this research, we employ two eye gaze trackers to learn the salient features in terms of eye movement data utilized by human participants for performing face age estimation and age-separated face verification. The key contributions of this paper are listed below:

- To learn the cues utilized by humans for the tasks of face age estimation and age-separated face verification, eye gaze data from 50 participants is acquired using two eye gaze trackers. As per our knowledge, this is the first dataset with these unique characteristics.

- The first in-depth analysis of the eye gaze data is presented. We examine the facial cues employed by humans for estimating the age group of presented face images and verifying age-separated face images. We evaluate eight face parts including four local regions and four halves of the face stimuli to analyze their relevance in the two tasks using different eye gaze metrics.

2. Literature Review

The ability of humans to successfully predict the age of an individual based on the facial cues has prompted compelling research in this domain. In his review paper, Rhodes [1] concluded that humans are remarkably accurate in age estimation of unfamiliar faces and are robust towards transformations in faces. Recently, Porcheron et al. [16] demonstrated that facial contrast is a cross-cultural cue for age perception. Furthermore, humans are also capable of verifying age-separated images of an individual. Yadav et al. [4] evaluated human performance on age-separated face images and examined the significance of various facial regions on this task.

The evolution of eye gaze trackers has enabled us to observe the cognitive process undertaken by the human brain while performing a task. Due to this trait, eye gaze tracking is being successfully utilized in various studies. Henderson [17] noted the importance of eye movements due to the attentional system employed by humans during a visual task. Eye movements provide unobtrusive and real-time information of the ongoing visual and cognitive processing in the brain. In another study, Armann and Bülthoff [18] examined the eye movements of the participants for the task of identity morphing. Using the eye gaze data, they concluded that the viewing fixations are dependent on the difficulty of the task.

3. Data Collection

In the literature, there is a lack of any database related to facial aging which provides face stimuli and corresponding eye gaze data. Hence, an experimental protocol is designed to acquire the eye gaze information of human subjects while performing these tasks. The data collection protocol is described in more detail subsequently.

3.1. Participants

50 volunteers (24 males and 26 females) participated in the data collection with mean age = 22.71 years and standard deviation = 4.16 years. All the participants reported normal or corrected-to-normal vision. The study was approved by Institutional Review Board at West Virginia University.

3.2. Apparatus

Two eye gaze trackers: EyeTribe[1] ET1000 and GazePoint GP3[2] are utilized for recording the eye movements. Both EyeTribe and GazePoint trackers offer a sampling frequency of 60 Hz. Before data collection of each participant, calibration is performed using the typical nine-point calibration system. This is followed by validating the precision of the eye gaze trackers.

3.3. Stimuli

The stimuli face images utilized in this study are collected from the FG-Net Facial Aging Database [10] and IIIT-Delhi Facial Aging Database [19]. For evaluation purposes, 10 sets of stimuli are created and one set is randomly assigned to every participant. Each set contains the following two tasks:

1. The participants are presented with a face image and are asked to estimate the age group of the stimuli. The presented face image belongs to one of the following age groups: 0-5, 6-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, and >70. Sample images for this task are displayed in Figure 2.

2. Next, participants are shown a pair of age-separated face images and asked to determine if they belong to

[1] http://theeyetribe.com

[2] www.gazept.com/product/gazepoint-gp3-eye-tracker/

Figure 2: Sample stimuli face images shown to the participants for estimating the age of the stimuli. The presented stimulus belongs to one of the above-shown age groups.

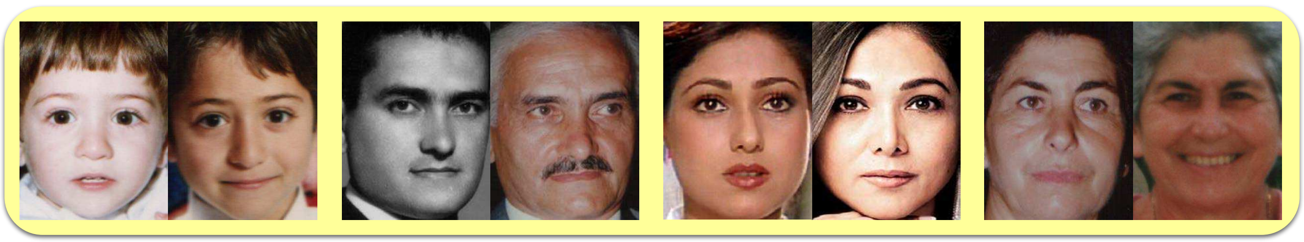


Figure 3: Sample stimuli face pairs presented to the participants for the task of age-separated face verification.

the same individual. Sample images for this task are displayed in Figure 3.

The stimuli are presented on a 24-inch screen with a resolution of 1920×1200 pixels.

3.4. Procedure

To acclimate participants to the eye gaze trackers, a training session is conducted on a per-participant basis. During the training phase, each participant is briefed about the two tasks and how they are supposed to provide their responses. Next, they are shown 10 sample stimuli face images and are asked to estimate their age groups from one of the 9 age group options. A similar process is followed for age-separated face pairs. After ensuring that the participant is familiar with the process, they are randomly assigned a set from one of the 10 assignment sets. This is followed by the nine-point calibration step described earlier.

Next, the participants are presented with the stimuli face images for age group estimation. The participants are instructed to select the appropriate age group of the presented stimuli. As soon as they indicate their response for a given image, the next image is shown. After completing the first task, the participants are allowed to take a break which is followed by re-calibration of the eye tracker. This is followed by the second task of verifying age-separated face images where pairs of age-separated stimuli face images are shown sequentially and the participants are asked to indicate if the shown pair of images belongs to the same person or not.

After acquiring eye gaze data from the EyeTribe tracker, the same process is repeated using GazePoint tracker. For both the trackers, the eye gaze data is stored as x-coordinate and y-coordinate of the point with respect to the display monitor.

4. Analysis of Eye Gaze Data

Using the earlier described procedure, eye movement data is captured with two gaze trackers from 50 users. In-depth analysis of the eye gaze data is performed to investigate the salient features in the stimuli face images for the tasks for age-group estimation and age-separated face verification. For analyzing the eye movements for different stimuli faces, the following metrics are computed to encode the eye gaze data:

- Number of fixations: A fixation is defined as the state when the visual gaze is focused on a single location. During a fixation, central foveal vision is held in place to allow the visual network to gather detailed information about the stimuli [13, 20]. The total number of fixations is the count of all the fixations by the participant on specific stimuli.
- Number of saccades: Saccades are quick eye movements which occur between different fixations [13, 21]. During saccades, the vision is restrained and new information is encoded only during the next fixation

Table 1: Accuracy (%) obtained by human participants for age group estimation of the presented stimuli face image.

Stimuli Age Group	Accuracy by Participants(%)
0-5	**82.45**
6-10	77.12
11-20	70.87
21-30	61.68
31-40	64.55
41-50	69.12
51-60	72.43
61-70	64.93
>70	**60.23**

state. The number of saccades is the count of different saccade eye movements across the stimuli.

- Gaze duration: It is the sum of all fixation durations in the given stimuli image. It is expressed in milliseconds (ms).

The analysis of responses and eye gaze is classified into four categories and key observations are discussed subsequently.

4.1. Age Group Estimation Accuracy

The behavioral responses provided by the participants for the first task of age group estimation are analyzed to determine their competence and the accuracy is computed. Accuracy is defined by the percentage of the number of correct responses as compared to the total number of responses. The results are tabulated in Table 1.

The overall accuracy in estimating the age group of the shown stimuli faces is 68.26%. Further analysis is performed and age group-wise accuracy is calculated. As seen in Table 1, stimuli face images from the age group 0-5 have the highest accuracy of 82.45% which is followed by images from the age group 6-10. A similar trend has been reported in the literature [4] which establish that face images belonging to newborns and toddlers are easiest to estimate.

On the other side of the spectrum, images belonging to the age group >70 are the most difficult to estimate with the lowest accuracy of 60.23%. Upon further analysis of the incorrectly predicted images of this age group, it is observed that most of the images are predicted incorrectly as 61-70 age group. This *neighboring bin effect* is also observed for the remaining age groups where the participants had incorrectly estimated the age group as the neighboring age group of the ground truth.

For comparative analysis of the responses provided by the participants, age group estimation of the same stimuli face images is performed using a commercial face recognition system: Face++[3]. Face++ yields 62.35% age group estimation accuracy with most errors in the age groups: 0-5, 6-10, 61-70, and >70. Sample age group estimation outputs are shown in Figure 4.

After analyzing the behavioral responses, we compute different eye movement based metrics to encode the eye

Participant	6-10	61-70	41-50
Face++	0-5	61-70	31-40
Ground Truth	6-10	61-70	31-40

Figure 4: Sample misclassification outputs by human participants and Face++ for age group estimation task.

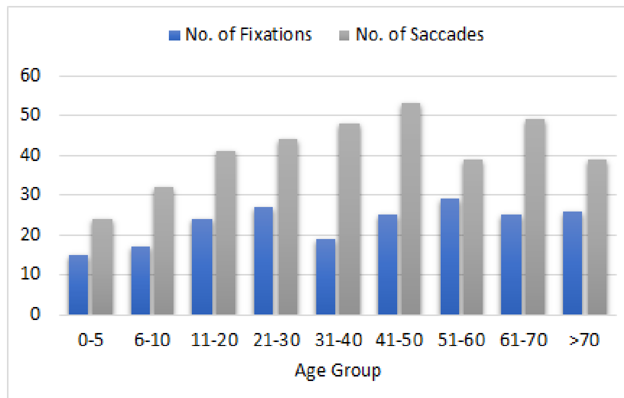


Figure 5: The average number of fixations and saccades across different participants for different age groups.

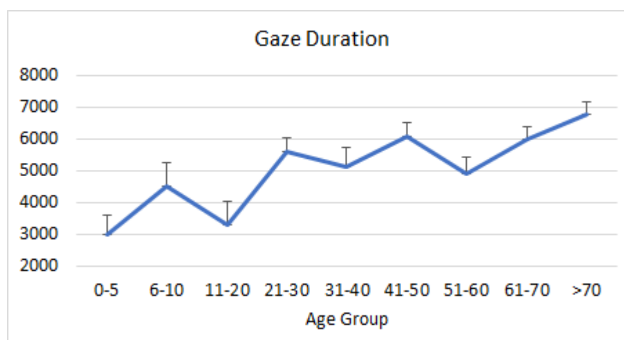


Figure 6: The average gaze duration for the age estimation task across different age groups.

[3] www.faceplusplus.com

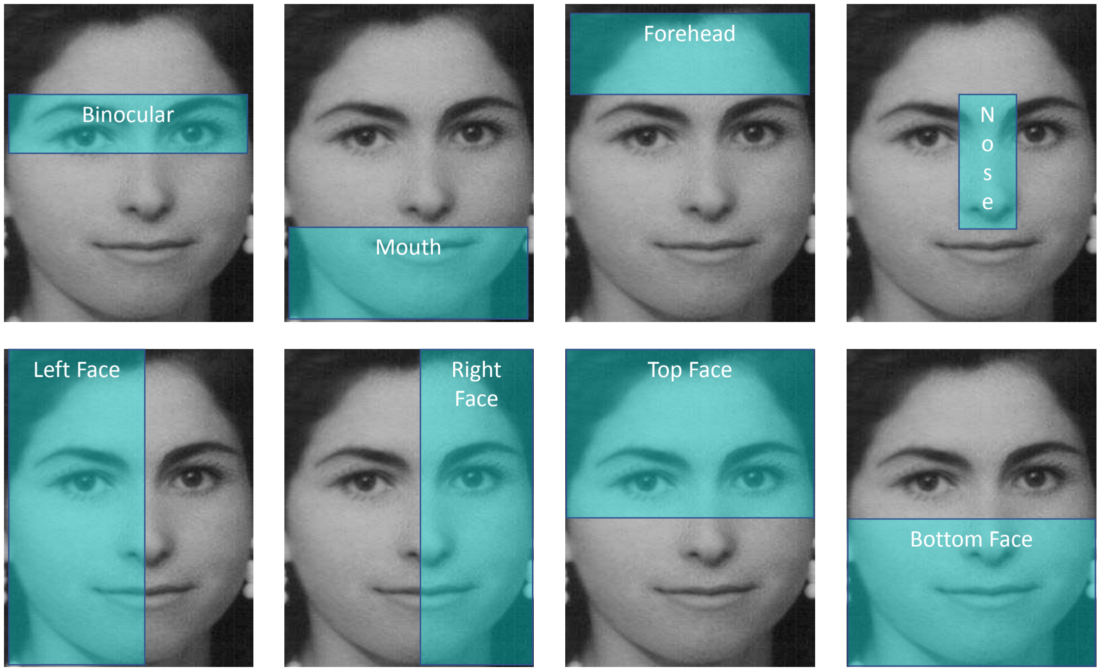


Figure 7: Selected areas of interest to examine the contribution of different facial regions.

gaze data of the participants while performing age estimation of stimuli faces. The different metrics computed across the different age groups as shown in Figures 5 and 6. With respect to the average number of fixations across a given face image for age estimation, it is observed that the age group 0-5 has the minimum number of fixations (15) while the age group 51-60 has the maximum number of fixations (29). In the literature, more number of fixations indicates less efficient search and difficulty in extracting information from the stimulus [13]. This indicates that participants spend less time on the stimuli from age group 0-5 and are able to quickly extract useful information for making the judgment. Similarly, upon comparing the average number of saccades across the different age groups, age group 0-5 has the minimum number of total saccades, i.e. 24. Even though no encoding takes places during saccades, less number of saccades indicates the less searching required for learning salient cues from the stimulus. An identical trend is observed for gaze duration as shown in Figure 6 with lowest gaze duration of 3009 ms for the age group 0-5.

4.2. Age-Separated Face Verification Accuracy

Next, the accuracy of human participants is evaluated on the second task, verification of age-separated face image pairs. It is observed that the participants correctly verified 71.19% age-separated face pairs. Upon further analysis, it is revealed that 69.23% of young face pairs, 73.45% of middle-aged, and 74.59% of old face pairs are correctly matched by the participants. The lowest verification accuracy of young stimuli face pairs has also been observed in the literature [4] and the drastic variations in the facial features during the childhood period is considered one of the possible reasons for the same.

Similar to the age group estimation task, the commercial face recognition system (Face++) is utilized to examine its face verification performance on the stimuli pairs shown to the participants. In the same experimental setup, Face++ yields 69.78% face verification accuracy which is 1.41% lower as compared to the human participants.

Analysis of the eye gaze data is performed and the aforementioned metrics are computed. The average number of fixation points are 41, 45, and 34 for young, middle-aged, and old stimuli face pairs respectively. The average of saccades across different stimuli face pairs are 37, 55, and 42 for young, middle-aged, and old stimuli face pairs respectively. The average gaze duration is 4502 ms, 5614 ms, and 4419 ms for young, middle-aged, and old stimuli face pairs respectively. Observing the values for the three metrics highlights that when the pair of stimuli face images belong to the old age range, participants spent the least time to accurately encode the features for matching these images.

4.3. Area of Interest-based Analysis

For analyzing the contribution of different facial regions and examining the salient cues, an area of interest based analysis is performed. The eye gaze data from the full face

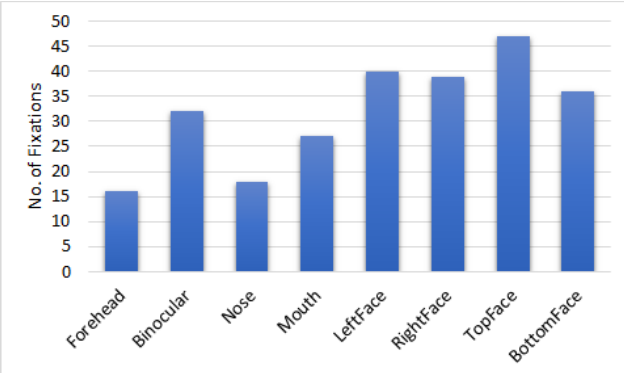


Figure 8: The average number of fixations across different areas of interest for the task of age group estimation.

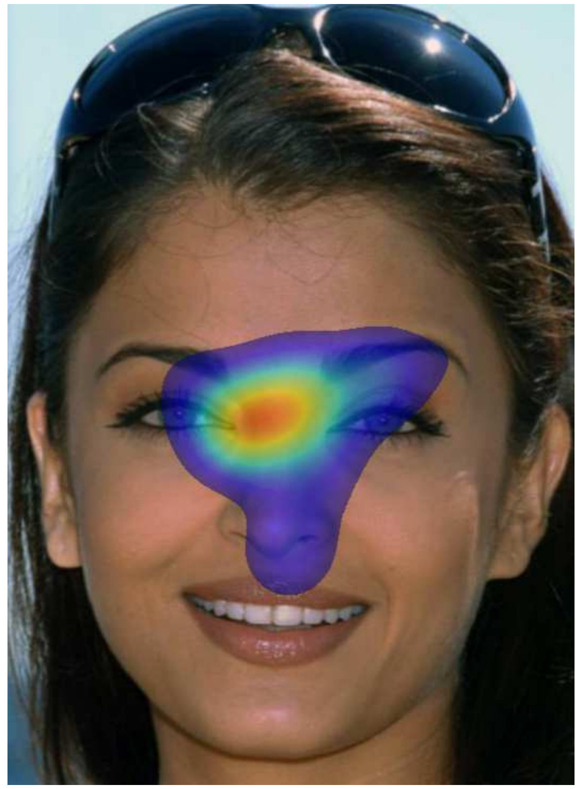


Figure 9: Sample heatmap demonstrating the number of fixations in different areas of interest. The blue-red gradient implies that the regions in red color have the maximum number of fixations while the regions in blue color have the minimum number of fixations.

is divided into the following regions (shown in Figure 7): (i) Binocular, (ii) Mouth, (iii) Forehead, (iv) Nose, (v) Left Face, (vi) Right Face, (vii) Top Face, and (viii) Bottom Face. These regions are selected based on different studies in the literature highlighting their significance in the task of facial aging [4, 1, 22].

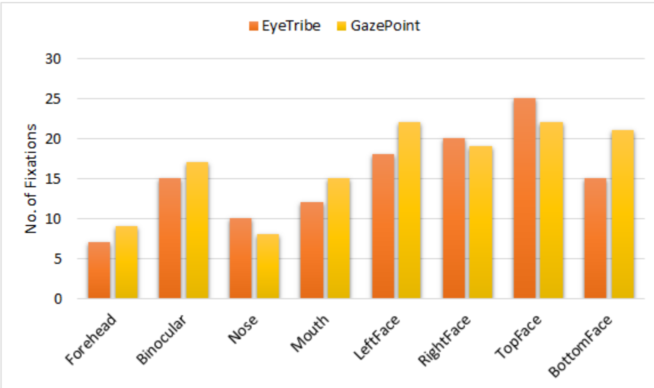
For the task of age group estimation using face images, the average number of fixations across the different areas of interest are plotted and shown in Figure 8. It is observed that among the four local facial regions, the binocular region containing the eyes is looked at the most number of times with the highest average number of fixations of 32. This indicates the existence of salient features in this region which are utilized by the participants for the age estimation task. This result is consistent with existing studies that highlight the significance of the eyes in predicting the facial age [5, 4]. On the other hand, the forehead region has the lowest number of fixations (16) which is half the number of fixations in the binocular region. The sample heat map shown in Figure 9 demonstrates the number of fixations across different areas of interest for age group estimation.

While comparing the different halves of the face area for age estimation, it is observed that the top half of the face gained more attention from the participants as compared to the lower half of the face. The average number of fixation in the upper face region is 47 which is 11 more than the number of fixations in the lower face region. These metrics highlight the areas of interest where the participants spent more fixations to encode the discriminatory features with respect to the different age group of the stimuli face image.

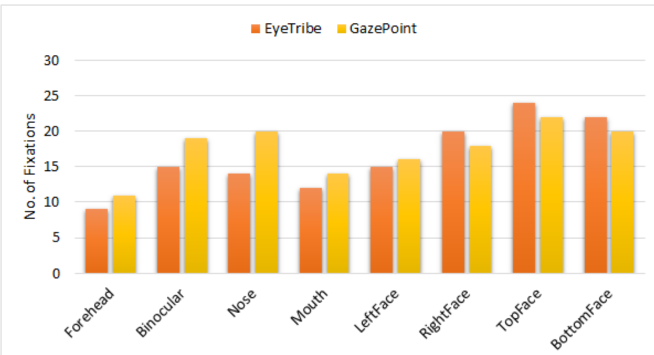
Next, the number of fixations across the eight areas of interest are evaluated for the second task, age-separated face verification. It is observed that among the four local regions, binocular region again has the highest number of fixations followed by nose, mouth, and forehead. This reveals that humans visit the binocular region (containing the eyes) for the maximum number of times to make this decision. Similar results have been shown for traditional face verification task, where the binocular region is the most frequently visited region [6, 23, 24]. Hsiao and Cottrell [7] showcased that while processing stimuli for face recognition, the initial fixations are more directed towards the center of the face stimulus, i.e. the nose area. Upon comparison of the different halves of the face, it is seen that the number of fixation points in the top half of the face is higher than the bottom half, indicating the presence of more relevant features in the upper face region.

4.4. Tracker-wise Analysis

A unique characteristic of the eye gaze data set collected in this paper is that it is acquired using two different eye gaze trackers: EyeTribe and GazePoint. As mentioned earlier, both the trackers operate at the sampling frequency of 60 Hz. In this subsection, we compare and contrast the data acquired from the two trackers with respect to the two tasks undertaken in this study. For this, the average number of fixations from EyeTribe and GazePoint across the eight areas of interest are compared. The results are summarized in



(a) Age Group Estimation



(b) Age-Separated Face Verification

Figure 10: Average number of fixations across different areas of interest from the two eye gaze trackers: EyeTribe and GazePoint.

Figure 10.

For the task of age estimation, it is seen that the forehead and nose regions have the least number of fixations for both the trackers. This illustrates that across both the trackers, the participants visited these regions least frequently to learn relevant features for this task. Upon comparing the salient features in left and right halves, the data from EyeTribe tracker illustrates a lower number of fixations in the left half as compared to the right half, unlike GazePoint tracker. Comparison of top and bottom halves of the face reveals congruous from both the trackers with top face containing more number of fixations as compared to the bottom half.

Figure 10 (b) shows the average number of fixations from EyeTribe and GazePoint for verifying age-separated face pairs. For both the trackers, forehead region has the least number of fixations across all the runs. This implies that the participants glanced at the forehead area least number of times to determine the similarity between the age-separated face pairs. Likewise, for both the trackers, the top half of the face images contains more number of fixations as compared to the bottom half.

5. Conclusion and Future Research Directions

Existing studies in the literature on human perception of facial aging are based on evaluating human performance using traditional questionnaire methodology. Responses to these questions may not provide an accurate representation of the cues utilized by the human participants. On the other hand, tracking of eye gaze movement can unveil knowledge about an individual's attention and corresponding cognitive activity for a specific visual stimulus. In this study, we utilize two eye gaze trackers to create first such database and to learn the salient features employed by human participants for the tasks of age group estimation and age-separated face verification. Area of interest analysis is performed to examine the significance of different local facial regions and it reveals the binocular area is a key region with the maximum number of fixations for age estimation. For age-separated face matching, the forehead region is visited the least, indicating the lack of salient features in that regions for the verification task.

In the future, we plan to extend the scope of this study to other covariates of face recognition such as disguise [25, 26] and plastic surgery [27, 28]. Further, we plan to utilize the knowledge learned from this study in designing novel algorithms, particularly in frame selection in video-based face recognition [29, 30].

Acknowledgments

The authors acknowledge the volunteers who participated in the data collection. M. Vatsa and R. Singh are partially supported by Infosys Center for Artificial Intelligence at IIIT-Delhi.

References

[1] M. G. Rhodes, "Age estimation of faces: A review," *Applied Cognitive Psychology*, vol. 23, no. 1, pp. 1–12, 2009. 1, 2, 6

[2] Y. H. Kwon and N. da Vitoria Lobo, "Age classification from facial images," *Elsevier Computer Vision and Image Understanding*, vol. 74, no. 1, pp. 1–21, 1999. 1

[3] W. Chen, W. Qian, G. Wu, W. Chen, B. Xian, X. Chen, Y. Cao, C. D. Green, F. Zhao, K. Tang *et al.*, "Three-dimensional human facial morphologies as robust aging markers," *Cell Research*, vol. 25, no. 5, p. 574, 2015. 1

[4] D. Yadav, R. Singh, M. Vatsa, and A. Noore, "Recognizing age-separated face images: Humans and machines," *PloS one*, vol. 9, no. 12, p. e112234, 2014. 1, 2, 4, 5, 6

[5] H. Han, C. Otto, and A. K. Jain, "Age estimation from face images: Human vs. machine performance," in *IEEE International Conference on Biometrics*, 2013, pp. 1–8. 1, 6

[6] M. F. Peterson and M. P. Eckstein, "Looking just below the eyes is optimal across face recognition tasks," *Proceedings*

of the National Academy of Sciences, vol. 109, no. 48, pp. E3314–E3323, 2012. 1, 6

[7] J. Hsiao and G. Cottrell, "Two fixations suffice in face recognition," *Psychological Science*, vol. 19, no. 10, pp. 998–1006, 2008. 1, 6

[8] M. Singh, S. Nagpal, R. Singh, and M. Vatsa, "On recognizing face images with weight and age variations," *IEEE Access*, vol. 2, pp. 822–830, 2014. 1

[9] L. Best-Rowden and A. K. Jain, "Longitudinal study of automatic face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 1, pp. 148–162, 2018. 1

[10] A. Lanitis, "Comparative evaluation of automatic age-progression methodologies," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, p. 101, 2008. 1, 2

[11] M. A. Just and P. A. Carpenter, "Eye fixations and cognitive processes," *Cognitive Psychology*, vol. 8, no. 4, pp. 441–480, 1976. 2

[12] Y. Li, P. Xu, D. Lagun, and V. Navalpakkam, "Towards measuring and inferring user interest from gaze," in *International Conference on World Wide Web Companion*, 2017, pp. 525–533. 2

[13] J. H. Goldberg and X. P. Kotval, "Computer interface evaluation using eye movements: methods and constructs," *International Journal of Industrial Ergonomics*, vol. 24, no. 6, pp. 631–645, 1999. 2, 3, 5

[14] M. Gingerich and C. Conati, "Constructing models of user and task characteristics from eye gaze data for user-adaptive information highlighting." in *AAAI*, 2015, pp. 1728–1734. 2

[15] R. A. Rensink, "The dynamic representation of scenes," *Visual Cognition*, vol. 7, no. 1-3, pp. 17–42, 2000. 2

[16] A. Porcheron, E. Mauger, F. Soppelsa, Y. Liu, L. Ge, O. Pascalis, R. Russell, and F. Morizot, "Facial contrast is a cross-cultural cue for perceiving age," *Frontiers in Psychology*, vol. 8, p. 1208, 2017. 2

[17] J. M. Henderson, "Human gaze control during real-world scene perception," *Trends in Cognitive Sciences*, vol. 7, no. 11, pp. 498–504, 2003. 2

[18] R. Armann and I. Bülthoff, "Gaze behavior in face comparison: The roles of sex, task, and symmetry," *Attention, Perception, & Psychophysics*, vol. 71, no. 5, pp. 1107–1126, 2009. 2

[19] D. Yadav, M. Vatsa, R. Singh, and M. Tistarelli, "Bacteria foraging fusion for face recognition across age progression," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2013, pp. 173–179. 2

[20] S. Marti, L. Bayet, and S. Dehaene, "Subjective report of eye fixations during serial search," *Consciousness and Cognition*, vol. 33, pp. 1–15, 2015. 3

[21] L. O. Wijdenes, L. Marshall, and P. M. Bays, "Evidence for optimal integration of visual feature representations across saccades," *Journal of Neuroscience*, vol. 35, no. 28, pp. 10 146–10 153, 2015. 3

[22] I. Huerta, C. Fernández, C. Segura, J. Hernando, and A. Prati, "A deep analysis on age estimation," *Pattern Recognition Letters*, vol. 68, pp. 239–249, 2015. 6

[23] M. Schurgin, J. Nelson, S. Iida, H. Ohira, J. Chiao, and S. Franconeri, "Eye movements during emotion recognition in faces," *Journal of vision*, vol. 14, no. 13, pp. 14–14, 2014. 6

[24] J. M. Henderson, R. Falk, S. Minut, F. C. Dyer, and S. Mahadevan, "Gaze control for face learning and recognition in humans and machines," in *From Fragments to Objects: Segmentation Processes in Vision*. Elsevier, 2000, pp. 1–14. 6

[25] T. I. Dhamecha, A. Nigam, R. Singh, and M. Vatsa, "Disguise detection and face recognition in visible and thermal spectrums," in *IEEE International Conference on Biometrics*, 2013, pp. 1–8. 7

[26] T. I. Dhamecha, R. Singh, M. Vatsa, and A. Kumar, "Recognizing disguised faces: Human and machine evaluation," *PloS one*, vol. 9, no. 7, p. e99212, 2014. 7

[27] R. Singh, M. Vatsa, H. S. Bhatt, S. Bharadwaj, A. Noore, and S. S. Nooreyezdan, "Plastic surgery: A new dimension to face recognition," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 441–448, 2010. 7

[28] N. Kohli, D. Yadav, and A. Noore, "Multiple projective dictionary learning to detect plastic surgery for face verification," *IEEE Access*, vol. 3, pp. 2572–2580, 2015. 7

[29] G. Goswami, R. Bhardwaj, R. Singh, and M. Vatsa, "MDLFace: Memorability augmented deep learning for video face recognition," in *IEEE International Joint Conference on Biometrics*, 2014, pp. 1–7. 7

[30] G. Goswami, M. Vatsa, and R. Singh, "Face verification via learned representation on feature-rich video frames," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 7, pp. 1686–1698, 2017. 7