

Feasibility Analysis of Eye Typing with a Standard Webcam

Yi Liu^{1,2}, Bu-Sung Lee², Andrzej Sluzek³, Deepu Rajan², and Martin McKeown⁴

¹ Nanyang Institute of Technology in Health and Medicine, Singapore,
yliu028@e.ntu.edu.sg

² SCSE, Nanyang Technological University, Singapore

³ Khalifa University, UAE

⁴ University of British Columbia, Vancouver, BC, Canada

Abstract. With the development of assistive technology, eye typing has become an alternative form of text entry for physically challenged people with severe motor disabilities. However, additional eye-tracking devices need to be used to track eye movements which is inconvenient in some cases. In this paper, we propose an appearance-based method to estimate the person's gaze point using a webcam, and also investigate some practical issues of the method. The experimental results demonstrate the feasibility of eye typing using the proposed method.

Keywords: assistive technology, eye typing, gaze estimation, appearance-based method

1 Introduction

In modern society, being healthy means more than just not being sick. To maintain good health, it is necessary to not only take care of the physical aspect but also the emotional one, with engagement with the community. Communication is an important tool to staying engaged. However, physically challenged people with motor disabilities have difficulty in communicating. About 120,000 cases of motor neuro disease (MND) are diagnosed worldwide every year [16]. People with these conditions may not be able to use speech, and movement of the eyes might be the only means by which they communicate.

In daily life, more than eighty percent of all information received comes from the eyes. Eyes and their movements play an important role in expressing human being's desires, cognitive processes, emotion states, and interpersonal relations. With advances in both eye-tracking technology and computer systems, eye-tracking research has gradually focused on human-computer interaction rather than just cognitive analysis. Eye-tracking devices have helped, in limited manner, some disabled people to enter text into their computer via eye typing system [15, 21, 25, 30], which provides an effective method of communication. However, early dwell-based eye-typing systems were limited in terms of the speed of text entry due to large dwell time [4, 14, 8, 19]. In order to increase the

speed, dwell-free eye-typing systems were proposed [9, 18, 17], but they are error-prone and cannot handle some common typing errors in practice [11]. Recently, Liu [31, 10] proposed a robust recognition method to address these challenging issues, even under low-accuracy eye tracker and low-quality calibration. However, the additional eye-tracking device (i.e. eye tracker) still needs to be used which is somehow burdensome. As the webcam is becoming a standard component in computers, especially in mobile devices, replacing the eye tracker with a webcam would simplify the setup of eye-typing equipment, and also facilitate the prevalence of eye-tracking applications.

Therefore, in this paper we investigated the feasibility of the development of eye typing system using a webcam as the input device. We showed that appearance-based method is an effective method for gaze estimation. During the course of our research we have improved the performance in terms of accuracy of gaze estimation by applying the average filter to the images. In addition, we investigated the effective eye areas that contribute the eye-appearance variance and finally determine the gaze estimation. According to the effective area analysis, the dimensionality of the eye-image feature can be reduced to achieve lower computational complexity. Finally performance evaluation indicates that the proposed method is sufficient to achieve a practical eye typing system using the standard system.

The reminder of the paper is organized as follows. Section 2 reviews related work. Section 3 describes the appearance-based method. Section 4 presents the results of the experiments followed by conclusion.

2 Related Work

Early research work on eye tracking with the standard webcam is mainly focusing on *eye detection/localization* [3, 6]. Eye detection deals with detecting the presence of eyes, accurately interpreting eye positions in the images, or tracking the eyes from frame to frame for video images, while eye localization refers to individual eye areas, i.e. to localize eyebrows, eyelids, eye corners, sclera, iris, and pupils. Various established techniques [29, 33, 34, 7, 23, 26] have been proposed to detect/localize eyes, in which most of them are robust in both indoor and outdoor environment, under different eye poses, and even with some degree of occlusion. More recently, the detected eyes in images or videos are used to estimate and track what a person is looking at, and this process is called *gaze estimation* [5]. Gaze estimation can be either a gaze direction in the 3D space, or a gaze point which is the intersection of gaze direction and a 2D plane.

In general, a person's gaze is determined by head pose and eyeball orientation as shown in figure 1. Usually head movement and eyeball rotation occur simultaneously, in which the person moves the head to a comfortable position before rotating the eyeball. Therefore, both head pose and pupil position need to be modelled for gaze estimation, which is very challenging. Currently most of research work on gaze estimation makes an assumption that the head pose is fixed, and only considers the eye area [24, 12, 13]. Thus the main task of gaze

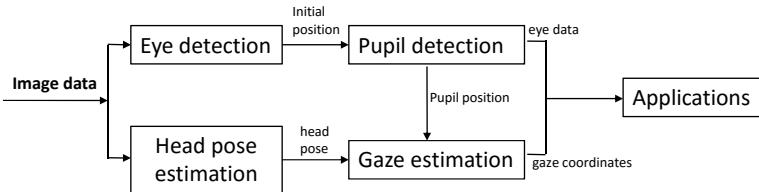


Fig. 1: Gaze determination with pupil position and head pose

estimation is modelling the relation between the image data of eye areas and the gaze direction/point. Similarly, we have adopted this same assumption in our gaze estimation study. Basically there are three types of approaches on gaze estimation [5], 3D-model-based, feature-based and appearance-based approaches.

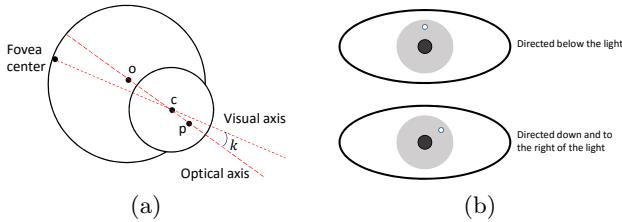


Fig. 2: The 3D-mode-based and feature-based approaches. (a) shows the 3D eye structure. The optical axis is connecting the pupil center, cornea center, and the eyeball center. The visual axis is connecting fovea center and cornea center, which is the true gaze direction. The angular offset between visual axis and optical axis is called *kappa* (k), which is a person-dependent constant. (b) shows the relation between pupil center and the cornea reflection

The 3D-model-based approaches [2, 28, 27] try to explicitly model the visual dynamics of the eyeball, and a simplified structure of the eyeball is shown in figure 2a. Although the visual axis is the actual gaze direction, the optical axis is usually determined instead, because the eyeball center and pupil center are relatively easier to estimate, and the angle between visual axis and optical axis is constant to each person. In order to avoid explicitly calculating the intersection of gaze direction and the 2D plane, the feature-based methods assume the underlying mapping from the eye features (e.g. pupil, iris, eye corners) to the gaze coordinates. Pupil center-cornea reflection vector [1] is the most common method to estimate the gaze point shown as figure 2b. Due to the limitation of the infrared light, some research work also suggests that the pupil center-eye corners vector can be an alternative which has the acceptable accuracy [22].

However, the 3D-model-based and feature-based approaches require highly-accurate feature detection, and are prone to errors. In addition, they usually need a high-resolution camera and infrared light. Unlike the above two approaches,

appearance-based approaches use the whole image content as an input that maps to gaze coordinates without the explicit local feature extraction [24, 12, 13]. In addition, the setup is more flexible, and the single webcam with relative low resolution is sufficient. Thus the appearance-based method is becoming promising gaze estimation technique for gaze estimation.

In the paper, we further investigated the performance of approach-based method in gaze estimation. We also discussed some practical issues of the method. Finally we investigated the feasibility of eye typing using the method. Although the accuracy is still less than the commercial eye tracker, it demonstrates the potential of eye typing using a standard webcam.

3 Appearance-based Method

Instead of extracting local features of eyes, appearance-based methods use an entire image of the eye as a high-dimensional input. The image is described as a feature vector in a high-dimensional space. A number of feature vectors will constitute a manifold which has an approximately 2D surface, because the eyeball movement has only 2 degrees of freedom. Assuming locally linear combination existing in the manifold, the 2D gaze points are estimated with the same locally linear mapping.

3.1 Eye-Appearance Feature Vector

The first task in an appearance-based approach is to crop the whole image and extract eye images (containing the eye area) as shown in figure 3. The RGB images are first converted to gray-scale images. Haar-cascade classifier is used to extract the rough eye regions in the image. The extracted images usually contain the eyebrow, which is removed using simple integrated projection technique. The eyebrow removal technique basically integrate the intensity values across a row of pixels. This will generate two global peaks (darkest parts), and using a threshold to eliminate the area around the upper peak. Canny edge filter is used to detect the inner and outer corners [12]. The eye image is cropped with a fixed size (figure 3a) in terms of the corners. We initially set the aspect ration to 60×36 pixels as suggested in [32]. Finally the feature vector of the eye appearance is generated by raster scan of the eye intensity images (figure 3b), and each pixel in the original image corresponds to an element in the vector. The eye-appearance image can be regarded as a point in a high dimensional space (i.e. the 2160-dimensional space). A set of these points will constitute a manifold in the high dimensional space, called eye-appearance manifold.

3.2 Eye-Appearance Manifold

Although the eye-appearance manifold is being in a high dimensional space, it has an approximately 2D surface [12], because the eyeball rotates in the 3D space. Therefore, when a person looks at different keys on an on-screen virtual

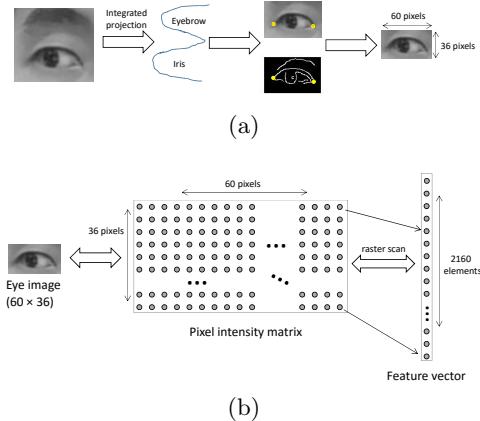


Fig. 3: Eye-appearance feature generation. (a) shows the procedures of cropping eye image. (b) illustrates the feature vector generation from the cropped eye image.

keyboard, the eyeball rotations are different, and thus the corresponding eye-appearance points are differentiable in the manifold. To verify this, we conducted a preliminary experiment. A person sat in front of a monitor with a webcam, and he was asked to look at each character key in order on the virtual keyboard (figure 4a) which was used in [10, 31]. When he looked at a key, fifty consecutive images of his frontal face were captured by the camera. After generating the eye-appearance feature vectors, we use PCA transformation to project these eye-appearance points into 2D/3D space as shown in figure 4.

There are some interesting observations: 1). The first three components contain most of the information (around 90% of accumulated eigenvalues), which is consistent with the observation in [12]. 2). The eye-appearance points can be clearly separated in 2D space. In another view, we can regard this as a classification of the 26 keys into 26 classes. 3). The relative positions of gaze points in 2D coordinate are maintained in the eye-appearance manifold. 4). The eye-appearance manifold is approximately a 2D surface embedded in a high dimensional space, and then it is possible to be modelled as a regression problem.

These observations help us understand the intuition of the appearance-based method, and how the eye appearance in high dimensional space changes with various gaze points in 2D space. Although the appearance-based methods do not estimate explicit parameters, (e.g. the eyeball centre, pupil centre and kappa in 3D-model-based methods, and pupil center-cornea reflection vector in feature-based methods), some underlying key parameters that determine the mapping relation between the eye appearance points and gaze points could be obtained implicitly.

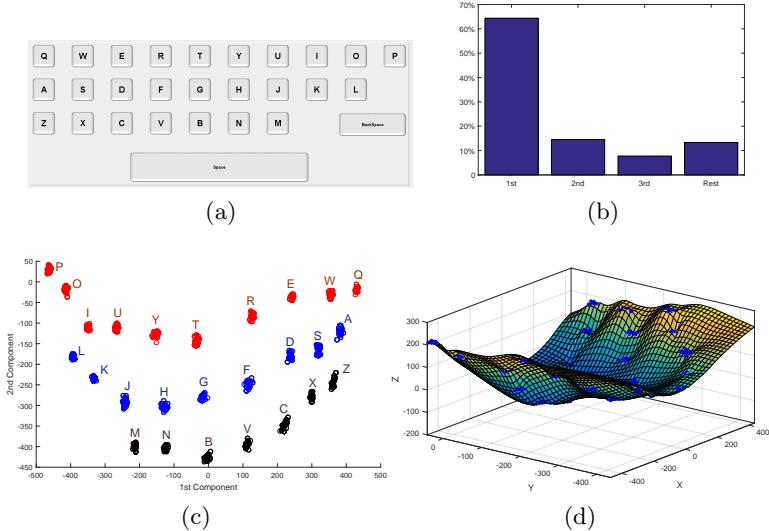


Fig. 4: The PCA projection. (a) shows the 26-letter keyboard layout used as the reference points. (b) shows the percentage of eigenvalues. (c) shows the first two principal components transformed by PCA into 2D space with the corresponding key labels. (d) illustrates the first three principal component forming the 2D surface which is the eye-appearance manifold in 3D space.

3.3 Locally Linear Embedding

In order to maintain the relative positions of gaze points in the eye-appearance manifold, the interpolation-based method with locally linear embedding [20, 24] is used to find the mapping relation. The method finds locally linear mapping instead of global mapping directly from high-dimensional data to low-dimensional data. The basic idea is estimating the mapping parameters (weights) of a new high-dimensional data point from its neighbouring data, and assuming the low-dimensional space also shares the same locally linear mapping, and then applying the same parameters to the corresponding low-dimensional data points to obtain the new low-dimensional data point (eq.1).

$$\begin{aligned} Xw &= \hat{x} \\ Pw &= \hat{p} \end{aligned} \tag{1}$$

where X is a matrix consisting of eye-appearance feature vector, P is the matrix consisting of corresponding gaze points, and w is the parameter that needs to be estimated. We do not estimate the mapping parameters between X and P . Given a new eye-image data point \hat{x} , we find the local mapping parameter w among its neighbors, and then estimate the gaze point \hat{p} , by assuming it shares the same neighboring mapping with the new eye-image data. However, it contains errors between the locally linear combination Xw and the new eye-image point \hat{x} , because the equation is overdetermined. The objective is to minimize the error

by tuning parameter w . Therefore, we have the gaze estimation function with optimization constraints in eq.2

$$\begin{aligned}\tilde{w} &= \arg \min \left| \hat{x} - \sum_i^k w_i x_i \right|, \text{s.t. } \sum_i^k w_i = 1 \\ \hat{p} &= \sum_i^k w_i p_i\end{aligned}\quad (2)$$

where \tilde{w} is optimal weight vector formed by the scalars w_i . $p_1 \dots p_k$ are the corresponding gaze points in 2D space. Once the optimal weights \tilde{w} are obtained, the gaze point \hat{p} is estimated using the same weights.

4 Experiments

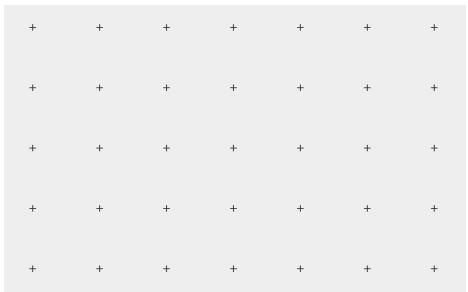
To further verify our observation and deepen our understanding of eye gaze estimation, we designed a more in-depth experiment. Other related work [12, 13] has already investigated the effect of the different number of reference points. In our experiments, our objectives were two fold: 1). improving the estimation of gaze position by having better calibration reference points. 2). determining the important features of image of the eye in determining the gaze point. This is done by reducing the size of the image and thus reducing the dimension size and complexity. In addition, we carried out some initial comparative study of eye typing using the appearance-based method and using the eye-tracker system.

4.1 Data Collection

We developed the system on a desktop computer with a 23-inch LED-lit monitor attaching an off-the-shelf webcam (30fps) as shown in figure 5a. The eye tracker (TheEyeTribe, 30Hz) was placed under the monitor. The chin rest, placed 50 cm in front of the monitor, is used to minimize the head movement of the participants. The experiment procedure is as follows: the participants first did the nine-point calibration with the eye tracker. Thirty-five cross-hair markers are individually displayed on the screen and the participants are asked to look at the central of the marker, and the participants were asked to look at the marker centre. To help the participants fixate at the centre, they were instructed to move the mouse cursor to align with the marker, and shape of the mouse cursor is the same as the marker [24]. When the mouse cursor overlaps with the marker, the participants are required to click on the mouse key. The webcam will simultaneously capture 10 images of the participants. Both the marker's position and the gaze coordinate estimation of the eye tracker are also recorded. Three volunteers (3 male, 25-33 years) from the local university participated in the experiment. All of them had normal vision. One of the participants had prior experience with the use of eye tracking input software/devices while the remaining two volunteers are novice.



(a)



(b)

Fig. 5: Experiment setup. (a) shows the placement of the monitor, webcam, eye tracker, chin rest. (3) shows the layout of reference points

The leave-one-out cross validation is employed in which one eye image is used as the test image and the rest for training. To measure the accuracy, the mean estimated angular error is calculated as eq.3:

$$\text{error} = \frac{1}{n} \sum_{i=1}^n \arctan\left(\frac{\|\hat{p}_i - p_i\|_2}{d}\right) \quad (3)$$

where $\|\hat{p}_i - p_i\|$ is the Euclidean distance between estimated gaze position \hat{p}_i and actual gaze position p_i . d is the distance between participant's eyes and the screen. We assume that the line of sight of the participant is perpendicular to the screen when the participant fixates at the centre.

4.2 Artifacts Elimination

In this experiment, we investigate the effect of image bias caused by artifacts using the appearance-based method. Since the reference points are very important for estimating the new gaze point, small bias of the reference points might result in large error in subsequent gaze estimation. In the experiment of previous work [24, 12], only one image of the participant's appearance was captured while doing the calibration. However, although the experiment is in a well controlled laboratory, some uncontrollable factors could affect the captured image, such as fine motion of eyeball, instant illumination variance. These artifacts will cause some bias of eye-appearance images even though the person is fixating at the same point on the screen, resulting in large error of gaze estimation. Therefore, it is necessary to eliminate or at least reduce/minimize its effect during eye gaze calibration period.

In our experiment, multiple images (10 images) were captured once the participant clicked the reference point while doing the calibration. Since 35 reference points are used for calibration, there are 35 sets of images where each set contains 10 images (figure 6a). If we use only one image from each set of images, there are 10^{35} images combinations, and we randomly selected only 10,000 combinations,

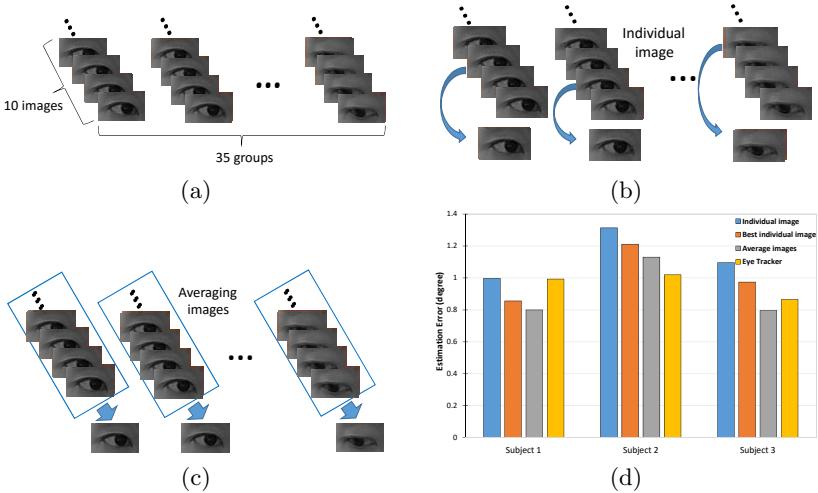


Fig. 6: Comparison of selecting individual image and averaging images

i.e. randomly selected one image from each group respectively (figure 6b), and repeated ten thousand times. For minimizing the effect of artifacts, we used an average filter to 10 images of each group (figure 6c). Figure 6d shows the result of the comparison. There are three groups of bars indicating three participants. In each group, the first bar is the average error of using individual image (average error of 10,000 combinations), and the second bar is the minimum error of using individual image (minimum error of 10,000 combinations). The third bar is the error using the proposed average filter. The fourth bar is gaze estimation error of the eye tracker. From figure 6d, the gaze estimation errors using the proposed average filter is 0.2 degree less than using individual images (average of 10,000 images), and 0.1 degree when compared with the minimum error when using individual image.

The results confirm our initial observation that artifacts of images captured in the calibration process could greatly affect the gaze estimation, and our proposed average filter is a good method to reduce the error. We plot the respective manifolds into 3D space by PCA transformation for visualization. Figure 7a shows the projection of all eye-image points, and figure 7b shows the projection of average eye-image points. As we can see, the manifold using the average filter has a much smoother surface, and thus it has a better locally linear mapping to estimate the gaze point.

4.3 Effective Area Detection

In this experiment, we investigate the effect of the size of the eye area on gaze estimation. First we define three cropping operations, horizontal cropping, vertical cropping, and full cropping (figure 8a). A single horizontal cropping operation is cropping a row of both top and bottom pixels of the eye image. A single vertical

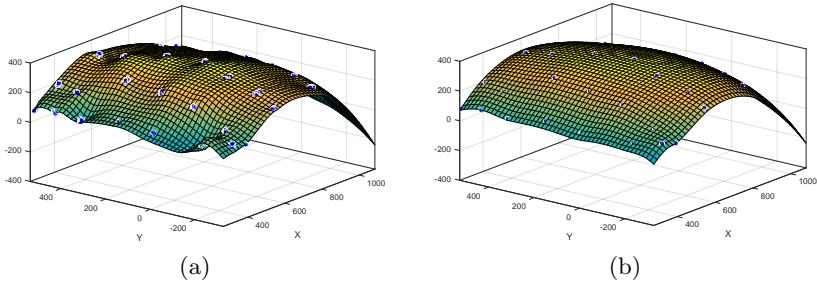


Fig. 7: Eye-appearance manifolds in 3D space. (a) shows the manifold consisting of all individual images. (b) shows the manifold consisting of the average images.

cropping operation is cropping one column of both right and left pixels. A single full cropping operation is a combination of horizontal and vertical cropping. The original size of the eye image is 60×36 pixels. Figure 8b, 8c, 8d show the results of the three cropping operations, where the x-axis is the number of times that the image is cropped, and the y-axis denotes the estimated angular error. There are two observations: 1). the estimation errors remain almost constant as the image is cropped horizontally until the “knee” point, above which the estimating error increases drastically. It was observed that the “knee” point occurs when the upper eyelid is cropped; 2). the effect of vertical cropping is less than horizontal cropping. Although both outer and inner eye corners are cropped, the error does not increase drastically. This suggests that eye corners as used in feature based methods are less critical. If the image can contain the whole iris, the gaze estimation error does not increase drastically.

These observations help us understand which parts of the eye in the image that contribute more to eye-appearance variance. Besides the iris part, we find the the upper eyelid has large impact on the gaze estimation which is not the case for 3D-model-based and feature-based methods. In 3D-model-based and feature-based methods, they focus mainly on the movement of iris or pupil part. In appearance-based method, although the degree of eyelid openness does not determine the gaze direction, it changes the eye appearance of the person and then affect the gaze estimation. Meanwhile, these observations also help us to reduce the dimensionality of eye-appearance feature vector. For example, as shown in figure 8d, if we do 10-time full cropping, the performance is almost same as the using the original image, but the size is reduced from 60×36 pixels to 40×16 pixels, and the 2160 dimension is reduced to 640 dimension, which is much more efficient in time and computing complexity.

4.4 Eye-typing Experiment

In this experiment, we investigate the feasibility of eye typing using webcam. The setup of the monitor, webcam, eye tracker and chin rest is the same as our previous experiments. The reference points used in the calibration are the 26-letter keys. After calibration with the 26-letter keys, the subject was asked

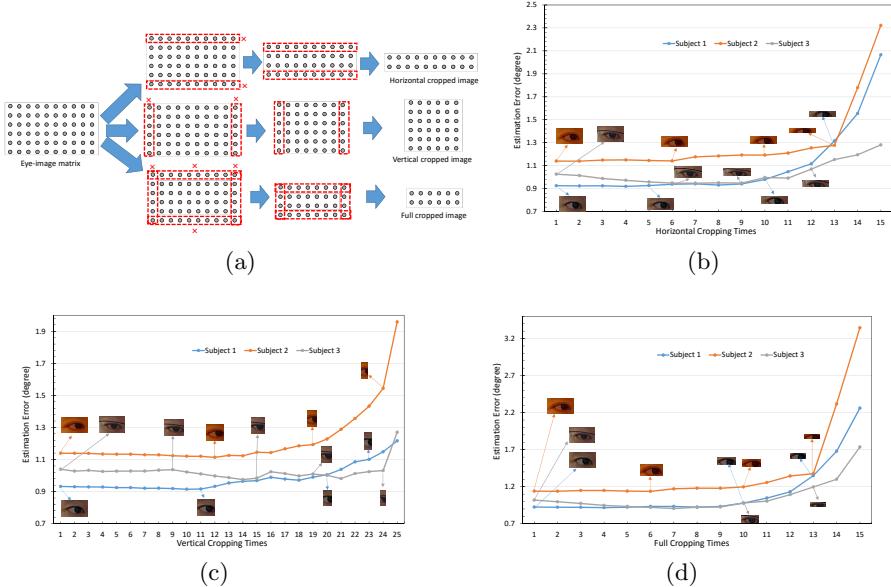


Fig. 8: Cropping operations. (a) shows the process of two-time horizontal, vertical, and full cropping operations, respectively. (b)(c)(d) show the error of horizontal, vertical, and full cropping operations in different cropping times. Different lines indicate different participants.

to type ten random words by gazing at letters of the words sequentially. The random-word selection refers to [10]. To help mark typing duration of each word, the participant clicked the mouse key alternatively (the first click for starting, the next click for stopping). The webcam recorded the video clips with 30fps, and estimated gaze points from the eye tracker (30fps, the same frequency with the webcam) were also recorded. All corresponding gaze points of these video clips were estimated using the proposed appearance-based method. In order to evaluate the performance of the method and the eye tracker, we used the recognition algorithm, LCSMapping, of the eye-typing system [10]. The LCSMapping algorithm will recommend top five words ranked by the probability based on the estimated gaze points, and if the intended word is listed in the top-5 words, it is regarded as a correct recommendation.

Figure 9 shows the gaze points on the keyboard coordinate estimated by the appearance-based method (blue dots) and the eye tracker (red dots) while typing the ten words respectively, and also shows the top-5 words by LCSMapping using estimated points of the appearance-based method (the upper right box) and the eye tracker (the lower right box). There are several observations, 1). The estimated points of eye tracker looks fewer, because eye tracker is more stable, and multiple estimated points have the same coordinates. These estimated points of eye tracker are more clustered together and accurate than the appearance-based method using the webcam, and the intended words are always listed at the

top-1 position; 2). although our appearance-based method is less accurate and there are small position shift, most of words are recognized and listed at top-1 position as shown in figure 9a-9f; 3). some words are still accurately recognized even though there is a large shift as shown in figure 9g-9h; 4). if the large shift causes the pattern of other words, the intended word cannot be recognized as shown in figure 9i-9j.

These observations help us understand the practical issues of appearance-based method. The calibration shift is the most critical factor that affects our appearance-based algorithm. The probable causes of the calibration shift are fine/small motion of the head and the degree of openness of eyelid. This will cause a higher number of both neighbour-letter and missing-letter errors in the eye-typing system [10]. The LCSMapping recognition algorithm can overcome some of these defects, however as the number of letter errors increase, the recognition algorithm begins to deteriorate. Therefore, it is feasible to achieve the eye typing using the webcam even with some degree of shift.

5 Conclusions

In order to improve the communication ability of people with motor disabilities, eye-based typing systems have been proposed. However, most typing systems require an external eye-tracking device. Recently, some eye-tracking research has been working on gaze estimation using the normal webcam. Compared with 3D-model-based and feature-based gaze estimation, the appearance-based method has higher potential of been used in the eye-typing area because of the simple setup without the high-resolution camera.

In the paper, we investigated whether the appearance-based method is able to differentiate eye-appearance points clearly while looking at different keys on the virtual keyboard, i.e. if it is feasible to classify the letter typed by eye movement. Our investigation found that the image bias caused by the uncontrollable artifacts existing in calibration would cause large error of gaze estimation. We proposed using an average filter to reduce the impact of the artifacts on gaze estimation. Another investigation is the determination of the effective area determining the gaze direction, it is possible to reduce the dimensionality of eye-appearance feature vector and decrease the time and space complexity. We also investigated the feasibility of eye typing using a webcam, in which we found that although the estimated points of the appearance-based method is less accurate than the eye tracker, the intended words can be recognized using the robust recognition algorithm in the eye-typing system.

In the future work, we will further investigate the impact of eyelid openness using the appearance-based method. The eyelid does not determine gaze direction, but it will affect gaze estimation. If we can find some specific patterns of eyelid and even reduce the impact, the accuracy of gaze estimation would improve. We will also integrate the appearance-based method with the robust eye-typing system, and more eye-typing experiments will be designed for investigation of practical issues of the method.

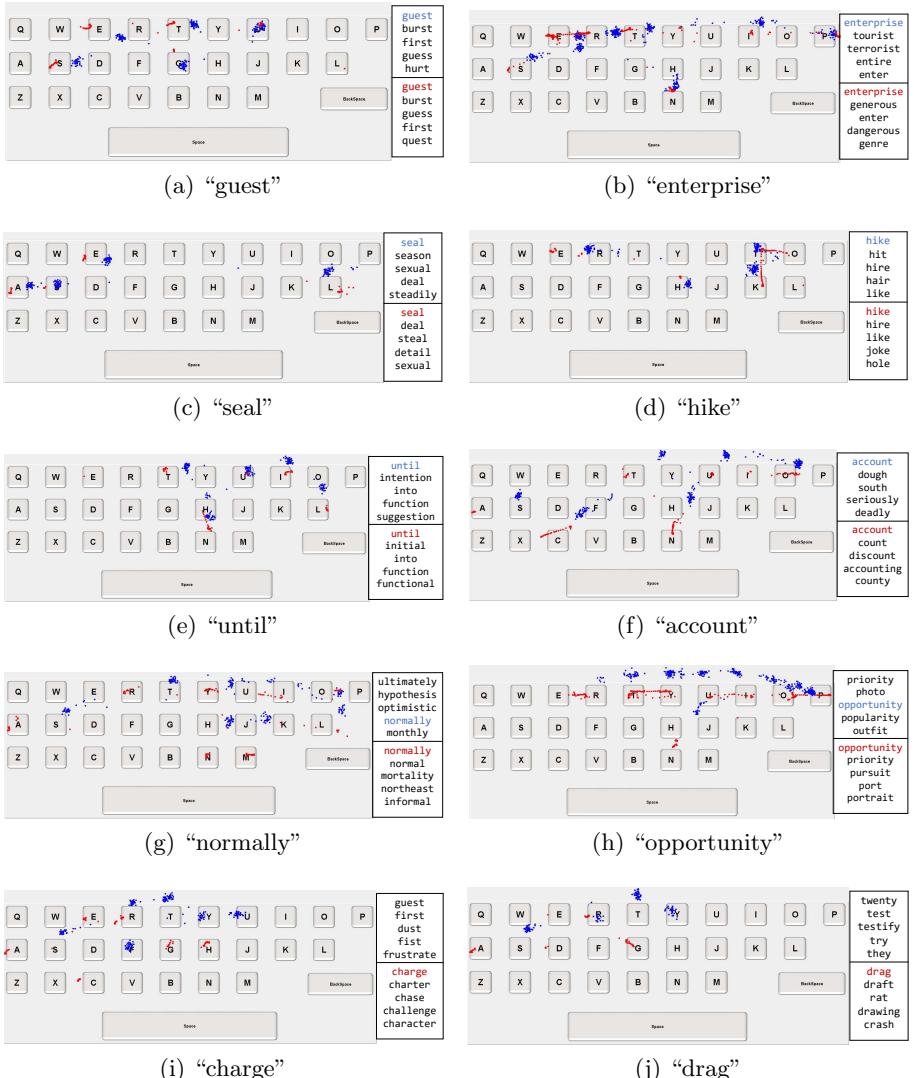


Fig. 9: Estimated gaze points of appearance-based method and eye tracker. They show the estimated points of typing ten words, “guest”, “enterprise”, “seal”, “hike”, “until”, “account”, “normally”, “opportunity”, “charge”, “drag”. The blue dots denote points estimated by the appearance-based method, and the red dots denote points by eye tracker. The upper right box shows top-5 words recommended using estimated points of the method, and the lower right box shows top-5 words using points of the eye tracker.

Acknowledgments. This work is a collaboration with the Joint NTU-UBC Research Centre of Excellence in Active Living for the Elderly (LILY).

References

1. Cerrolaza, J.J., Villanueva, A., Cabeza, R.: Taxonomic study of polynomial regressions applied to the calibration of video-oculographic systems. In: Proceedings of the 2008 symposium on Eye tracking research & applications. pp. 259–266. ACM (2008)
2. Chen, J., Ji, Q.: Probabilistic gaze estimation without active personal calibration. In: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. pp. 609–616. IEEE (2011)
3. Feng, G.C., Yuen, P.C.: Multi-cues eye detection on gray intensity image. Pattern recognition 34(5), 1033–1046 (2001)
4. Hansen, D.W., Hansen, J.P., Nielsen, M., Johansen, A.S., Stegmann, M.B.: Eye typing using markov and active appearance models. In: Applications of Computer Vision, 2002.(WACV 2002). Proceedings. Sixth IEEE Workshop on. pp. 132–136. IEEE (2002)
5. Hansen, D.W., Ji, Q.: In the eye of the beholder: A survey of models for eyes and gaze. Pattern Analysis and Machine Intelligence, IEEE Transactions on 32(3), 478–500 (2010)
6. Hansen, D.W., Pece, A.E.: Eye tracking in the wild. Computer Vision and Image Understanding 98(1), 155–181 (2005)
7. Kim, S.T., Choi, K.A., Shin, Y.G., Ko, S.J.: A novel iris center localization based on circle fitting using radially sampled features. In: Consumer Electronics (ISCE), 2015 IEEE International Symposium on. pp. 1–2. IEEE (2015)
8. Kotani, K., Yamaguchi, Y., Asao, T., Horii, K.: Design of eye-typing interface using saccadic latency of eye movement. International Journal of Human-Computer Interaction 26(4), 361–376 (2010)
9. Kristensson, P.O., Vertanen, K.: The potential of dwell-free eye-typing for fast assistive gaze communication. In: Proceedings of the Symposium on Eye Tracking Research and Applications. pp. 241–244. ACM (2012)
10. Liu, Y., Lee, B.S., McKeown, M.J.: Robust eye-based dwell-free typing. International Journal of Human-Computer Interaction (In Press) (2016)
11. Liu, Y., Zhang, C., Lee, C., Lee, B.S., Chen, A.Q.: Gazetry: Swipe text typing using gaze. In: Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction. pp. 192–196. ACM (2015)
12. Lu, F., Sugano, Y., Okabe, T., Sato, Y.: Inferring human gaze from appearance via adaptive linear regression. In: Computer Vision (ICCV), 2011 IEEE International Conference on. pp. 153–160. IEEE (2011)
13. Lu, F., Sugano, Y., Okabe, T., Sato, Y.: Adaptive linear regression for appearance-based gaze estimation. Pattern Analysis and Machine Intelligence, IEEE Transactions on 36(10), 2033–2046 (2014)
14. MacKenzie, I.S., Zhang, X.: Eye typing using word and letter prediction and a fixation algorithm. In: Proceedings of the 2008 symposium on Eye tracking research & applications. pp. 55–58. ACM (2008)
15. Majaranta, P., Ahola, U.K., Špakov, O.: Fast gaze typing with an adjustable dwell time. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 357–360. ACM (2009)
16. Majaranta, P., Aoki, H., Donegan, M., Hansen, D.W., Hansen, J.P.: Gaze Interaction and Applications of Eye Tracking: Advances in Assistive Technologies. Information Science Reference - Imprint of: IGI Publishing, Hershey, PA, 1st edn. (2011)

17. Pedrosa, D., Pimentel, M.d.G., Truong, K.N.: Filteryedping: A dwell-free eye typing technique. In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems. pp. 303–306. ACM (2015)
18. Pedrosa, D., Pimentel, M.D.G., Wright, A., Truong, K.N.: Filteryedping: Design challenges and user performance of dwell-free eye typing. ACM Transactions on Accessible Computing (TACCESS) 6(1), 3 (2015)
19. Räihä, K.J., Ovaska, S.: An exploratory study of eye typing fundamentals: dwell time, text entry rate, errors, and workload. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 3001–3010. ACM (2012)
20. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. *Science* 290(5500), 2323–2326 (2000)
21. Sarcar, S., Panwar, P., Chakraborty, T.: Eyek: an efficient dwell-free eye gaze-based text entry system. In: Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction. pp. 215–220. ACM (2013)
22. Sesma, L., Villanueva, A., Cabeza, R.: Evaluation of pupil center-eye corner vector for gaze estimation using a web cam. In: Proceedings of the Symposium on Eye Tracking Research and Applications. pp. 217–220. ETRA '12, ACM, New York, NY, USA (2012), <http://doi.acm.org/10.1145/2168556.2168598>
23. Skodras, E., Fakotakis, N.: Precise localization of eye centers in low resolution color images. *Image and Vision Computing* 36, 51–60 (2015)
24. Tan, K.H., Kriegman, D.J., Ahuja, N.: Appearance-based eye gaze estimation. In: Applications of Computer Vision, 2002.(WACV 2002). Proceedings. Sixth IEEE Workshop on. pp. 191–195. IEEE (2002)
25. Urbina, M.H., Huckauf, A.: Alternatives to single character entry and dwell time selection on eye typing. In: Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications. pp. 315–322. ACM (2010)
26. Valenti, R., Gevers, T.: Accurate eye center location through invariant isocentric patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 34(9), 1785–1798 (2012)
27. Villanueva, A., Cabeza, R., Porta, S.: Gaze tracking system model based on physical parameters. *International Journal of Pattern Recognition and Artificial Intelligence* 21(05), 855–877 (2007)
28. Wang, J.G., Sung, E., Venkateswarlu, R.: Estimating the eye gaze from one eye. *Computer Vision and Image Understanding* 98(1), 83–103 (2005)
29. Wang, P., Green, M.B., Ji, Q., Wayman, J.: Automatic eye detection and its validation. In: Computer Vision and Pattern Recognition-Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on. pp. 164–164. IEEE (2005)
30. Ward, D.J., MacKay, D.J.: Artificial intelligence: Fast hands-free writing by gaze direction. *Nature* (2002)
31. Yi Liu, Bu-Sung Lee, M.M., Lee, C.: A robust recognition approach in eye-based dwell-free typing. In: Proceedings of 2015 International Conference on Progress in Informatics and Computing. pp. 5–9. IEEE (2015)
32. Zhang, X., Sugano, Y., Fritz, M., Bulling, A.: Appearance-based gaze estimation in the wild. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 4511–4520 (2015)
33. Zhou, Z.H., Geng, X.: Projection functions for eye detection. *Pattern recognition* 37(5), 1049–1056 (2004)
34. Zhu, Z., Ji, Q.: Robust real-time eye detection and tracking under variable lighting conditions and various face orientations. *Computer Vision and Image Understanding* 98(1), 124–154 (2005)