# Anti-Masked Face Recognition and Eye Tracking for Direct-View Augmented Reality Surgical Navigation

**Jian Lung-Chen[1], Yu-Ju Chao[1], Ting-Hsun Cheng [1], Yu-Hsin Lin[1]**

Bigtim@itri.org.tw
[1]Industrial Technology Research Institute, Hsinchu, Taiwan 31040, R.O.C.

Keywords: Machine Learning; Facial recognition; Data augmentation; color space.

## ABSTRACT

*The world has been affected by the pneumonia (COVID-19), people began to wear masks in daily or working. It represented a huge challenge for face recognition. The algorithms would be less accurate if part of face was concealed in face detection and eye tracking in surgery room.*

*AR surgical system displayed surgeon's information about the lesion structure, and providing information in the surgery. The facial features are in partial occlusion. In this study, it proposed an anti-mask face recognition including eye tracking which can be implemented for direct-view AR in surgical navigation. This eye tracking technology combines a facial feature occlusion detection, and the ML algorithm to detect the eyes of the occluded face. The occluded features are repaired, and the machine learning library is used for facial landmark detection. Use the color space grouping technology to calculate the eyeball center point. The test verifies that the eyeball center accuracy has reached the level of 4 pixel.*

## 1 Introduction

The information provided by the classical surgical navigation system cannot be superimposed with the position of the lesion to guide the operation. This situation prevents the doctor from synchronize hand and eye, but must swing his head to catch the information displayed on the screen. When the doctor focuses on the surgical procedure, the shift of the visual focus in an instant can easily cause blurred vision and increase the risk of surgery. To solve this problem, we propose an anti-masking eye positioning algorithm, which can realize the fusion of virtual reality through virtual reality in surgical navigation with transparent display screen. Combining the adaptive obscured facial detection technology, facial image restoration technology, and facial landmark restoration technology, we achieved the purpose of eyeball center positioning. This technology with a transparent display is equipped to provide the surgeon surgical navigation assistance of hand and eye synchronization. Transparent display smart surgical navigation system, as shown in right side Figure 1, provided intuitive and burden-free surgical situation. Make Doctor Always on the target.

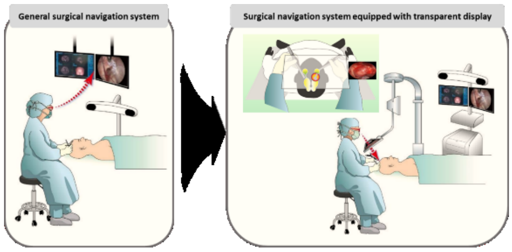


**Fig. 1 Evolution of surgical navigation system**

## 2 Experiment

Before the eyeball positioning for obscured face, the user's facial image must be captured and stored in database. The facial feature recognition algorithm is used for machine learning to perform facial feature detection and obtain the unmasked facial feature information position. Based on the direction and angle of head rotation, the most appropriate simulation model selection is performed. After the user's face image is restored, the position of the landmark near the eyes of the virtual image can be calculated based by a general cross-platform machine learning library. The self-developed HSV clustering algorithm is used to calculate the position of different areas (eyeballs, sclera, and skin) to achieve eyeball positioning, as shown in Figure 2.

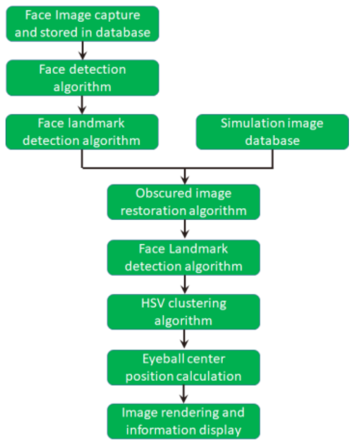


**Fig. 2 Eyeball positioning process for obscured face**

The procedure of anti-masked face recognition and eye tracking contains obscured face detection, facial landmark restoration, facial landmark positioning of restoration images, image processing and eye center position and tracking, has shown in Figure 3.

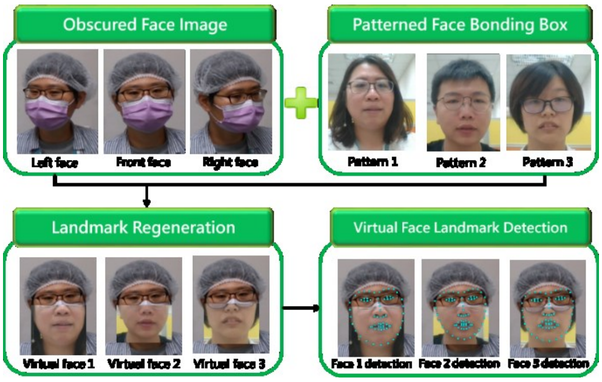


**Fig 3 Facial landmark restoration process for obscured face**

In order to achieve accurate eyeball positioning, the algorithm can be refined individually from the following points: 1) Consider the orientation of the user's face, and select the most appropriate non-masked face image in database to improve the detect accuracy of landmark position near eyeball by the general cross-platform machine learning library. 2) Revise the clustering algorithm for image near eyeball by adding attached blocks. With the addition of attached blocks, it is possible to more accurately and sharply distinguish the areas of the eyeballs, sclera, and skin. 3) We use AI inference to determine the reference eyeball center position. The test instructions are as follows:

### 2.1 User's face orientation

The user's face orientation can be determined by the following method. The aspect ratio of the bounding box and the image information of the bounding box are used to estimate the user’s face orientation. The relationship between the aspect ratio of the bounding box and the user’s face orientation has shown in Figure 4.

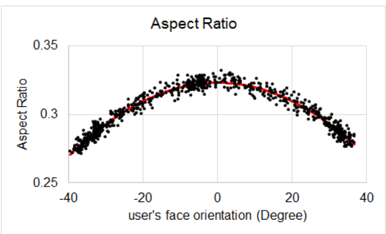


**Fig 4 Aspect ratio of bounding box versus user’s face orientation**

We can calculate the of head pose direction from bounding box’s aspect ratio by the following equation.

$$Aspect\ Ratio = 0.323 - \left(\frac{User's\ face\ orientation}{175}\right)^2$$

Or equivalently

$$User's\ face\ orientation = 175 \times \sqrt{0.323 - Aspect\ Ratio}$$

The direction the head is facing (left or right) can be determined by the sign of the direction indicator. If the direction indicator is positive, the head is facing right. Conversely, if the direction indicator is negative, the head is facing left. The direction indicator is calculated from the framed image, and the description is as follows:

The width of the bounding box is w pixels. First, the gray-scale signals of the bounding box are summed in the y direction, and then the differential of the sum in the x direction is calculated. Multiply the differentiated result with the following function g(x) and then add up to get the direction index. The flow chart of this procedure is shown in Figure 5.

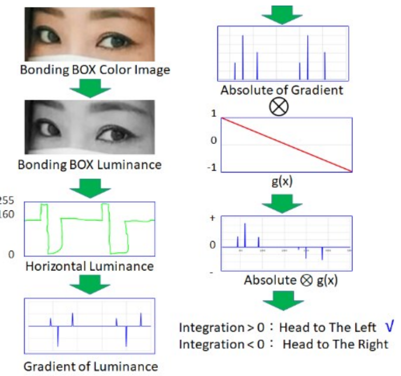


**Fig 5 Direction indicator calculation**

Figure 6 shows the results of head orientation and direction indicators. The correct rate of head orientation is 76.3%. Comprehensive head orientation angle, as shown in Figure 7 and Figure 8, the average judgment error of all the judgments is 5.19 degrees. Based on the above experimental results, we have built several bionic models with different head orientation angles in the database, from 40 degrees to the left to 40 degrees to the right, with an interval of 5 degrees.

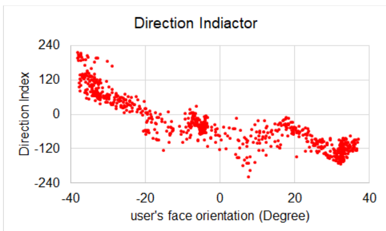


**Fig 6 Direction indicator vs user’s face orientation**

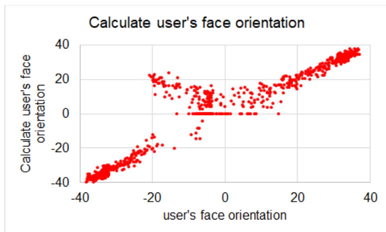


**Fig 7 Predicted and actual user’s face orientation**

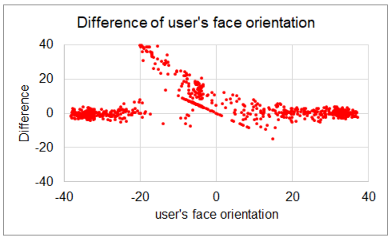


**Fig 8 Difference of predicted and actual orientation**

## 2.2 Revise the clustering algorithm

Revise the clustering algorithm for image near eyeball by adding attached blocks. With the addition of attached blocks, it is possible to more accurately and sharply distinguish the areas of the eyeballs, sclera, and skin. The instructions are as follows.

Set the image around the eyeball as the primary block, as shown in Figure 9.

Calculate the grayscale distribution of the primary block, pick up the brightest and darkest positions of the grayscale brightness, and record the RGB color level of the brightest and darkest positions.

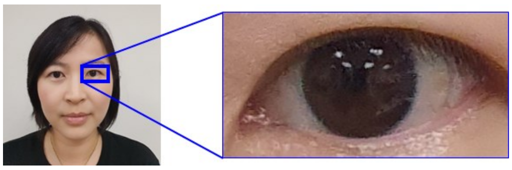

**Fig 9 Image near the eyeball center**

For example, as shown in Figure 10, the RGB color level of the brightest position is (blue: 37, green: 29, red: 40), and the RGB color level of the darkest position is (blue: 190, green: 198, red: 227).


**Fig 10 RGB color level of the brightest and Darkest position**

Add an attached block at the two sides of the primary block. Fill up the color of the "brightest" position of the primary block in the left attached block.

Fill up the color of the "darkest" position of the primary block in the left attached block, as shown in Figure 11.

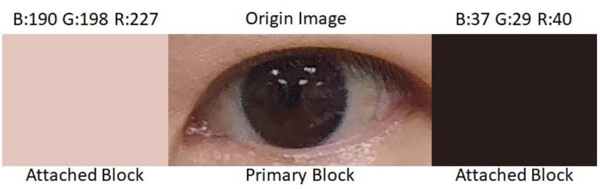


**Fig 11 Primary block and attached blocks**

Then the primary block and the attached blocks are combined to perform the self-developed HSV clustering algorithm, as shown in Figure 12. When the points in primary block and the right attached block are in the same cluster, it means that these points is in the eyeball position. Geometric center of these points is the center of the eyeball.

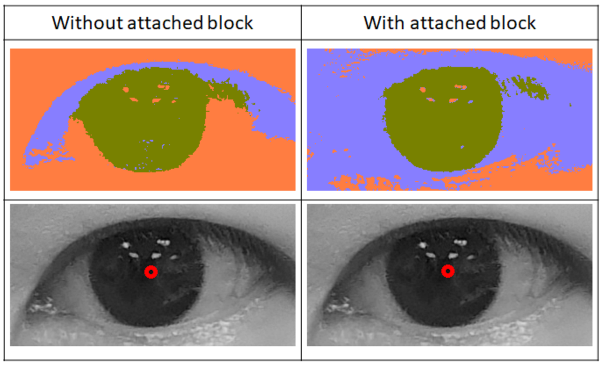


**Fig 12 Self-developed HSV clustering algorithm without and with attached blacks**

## 2.3 The reference eyeball center position

We use AI inference to determine the reference eyeball center position.

We train a Convolution Neuron Network model [1] using the labeled 1028 images of the left eyes and labeled 1028 images of the right eyes. The loss during training process is shown in Figure 13. The difference of the eyeball center between the labeled and predict has shown in Figure 14.

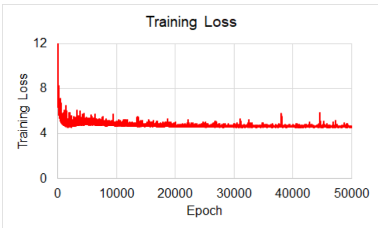


**Fig 13 Training loss of reference eyeball center**

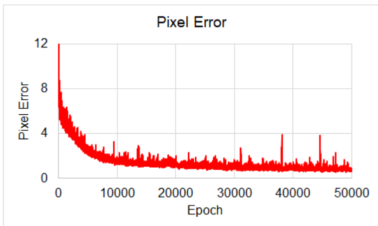


**Fig 14 Position difference between labeled and predict eyeball center**

## 3 Results

The distance between the eyeball position calculated by the revised self-developed clustering algorithm and the eyeball position predicted by the AI inference is the error of the revised self-developed clustering algorithm. a) As shown in Figure 12, the clustering result of the primary block will be affected by the size of attached block. b) When the ratios of single attached block and the primary block are different, as shown in Figure 15, the calculation results of the eyeball center position are also different, as shown in Figure 16. c) The error of eyeball center position of different ratios is shown in Figure 17.


**Fig 15 Different block ratios a) single attached block area : primary block area=1:2 , b) single attached block area : primary block area=1:1**

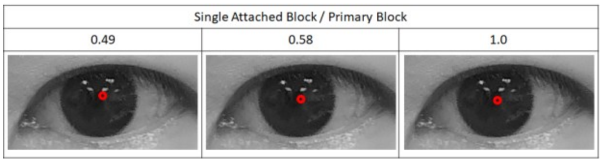


**Fig 16 Eyeball center position versus block ratio**

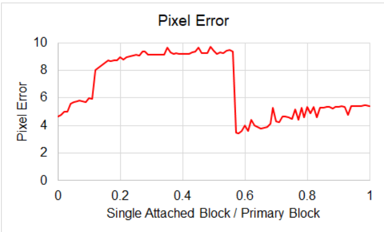


**Fig 17 The eyeball center error versus block ratio**

## 4 Discussion

From the above test, the aspect ratio of the bounding frame and the image signal of the bounding frame can be used to estimate the orientation of the user's face, and from the database, select the most appropriate image to restoration landmarks of the obscured face. We verify the improved clustering algorithm results with the AI inference result. When the ratio of a single attached block to the primary block is 0.58, the best accuracy of the eyeball center position is achieved.

## 5 Conclusions

This research integrates obscured face detection and an eyeball positioning algorithm based on image clustering, developed an eyeball positioning mechanism for partial face occlusion. It can be applied to precise image fusion in the surge. Through improvements and tests, the accuracy of the eyeball center position can reach 4 pixels.

## References

[1] Liu, W., Dragomir Anguelov, D. Erhan, Christian Szegedy, Scott E. Reed, Cheng-Yang Fu and Alexander C. Berg. "SSD: Single Shot MultiBox Detector." ECCV (2016).

[2] Redmon, J. and Farhadi, A. (2018) YOLOv3: An Incremental Improvement. Computer Science, arXiv: 1804.02767. http://arxiv.org/abs/1804.02767

[3] Jin X., Han J. (2011) K-Means Clustering. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-30164-8_425

[4] Davis E. King. Dlib-ml: A Machine Learning Toolkit. Journal of Machine Learning Research 10, pp. 1755-1758, 2009

[5] (2008) HSV Color Space. In: Li D. (eds) Encyclopedia of Microfluidics and Nanofluidics. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-48998-8_656

[6] Podpora, Michal. (2009). YUV VS. RGB--A COMPARISON OF LOSSY COMPRESSIONS FOR HUMAN-ORIENTED MAN-MACHINE INTERFACES. Zeszyty Naukowe. Elektryka. 55-56.

[7] Szeliski R., 'Computer Vision: Algorithms and Applications', Springer, 2011.

[8] Hunter, Richard Sewall (July 1948). "Photoelectric Color-Difference Meter". JOSA. 38 (7): 661. (Proceedings of the Winter Meeting of the Optical Society of America).

[9] Hunter, Richard Sewall (December 1948). "Accuracy, Precision, and Stability of New Photo-electric Color-Difference Meter". JOSA. 38 (12): 1094. (Proceedings of the Thirty-Third Annual Meeting of the Optical Society of America).

[10] Mezher, Alhamzawi Hussein Ali. "Real Time Gaze Estimation for Medical Field using Normally Webcam with OpenCV." International journal of scientific research in science, engineering and technology 3 (2017): 136-141.

[11] Ahmed, Noha. (2020). Real-time accurate eye center localization for low-resolution grayscale images. Journal of Real-Time Image Processing. 10.1007/s11554-020-00955-2

[12] Bradski, G. (2000) The OpenCV Library. Dr. Dobb's Journal of Software Tools.