Software and Hardware Implementation of DMS Based on Infrared Camera Behind the LCD

Yating Wen [1], Xiuyan Li [2], Yufeng Jin, Yu Wu, Chris Kim, Tao He, Keming Chen, Chao Tian, Fei Ai, Qiang Gong, Jingfeng Xue, Bin Zhao, Yuanmin Deng, Mulin Liu, Fan Zhang, Song Yang

[1] Shenzhen TCL China Star Optoelectronics Technology Co., Ltd, Guangdong, China

[2] Wuhan TCL China Star Optoelectronics Technology Co., Ltd, Hubei, China

Abstract

Driver monitor system (DMS) is an essential configuration for safe driving. In order to miniaturize the car frame and maximize the interior space, we integrated the infrared camera behind the liquid crystal display (LCD). Unlike typical blind hole screens, we adopted a full-screen design with the same pixel architecture throughout. Its backlight has no light-transmitting holes and is a complete infrared high-transmittance film. When the infrared transmittance of the LCD full screen is 30%, the camera can receive a blurred image with contours. To solve the problem of image blurriness, we consider the LCD full screen as composed of thin films of different media. The image becomes increasingly blurred with each layer of film. One layer of film is similar to a layer of fog. After layer-by-layer guided filtering and de-fogging, the image becomes clear. After multi-layer defogging, the facial recognition ratio (FRR) has increased by nearly 60% compared to before. In addition, we propose two parameters representing abnormal head states based on two-dimensional facial images: HTR and HTG, which can sensitively reflect head turning and tilting actions.

Author Keywords

Driver monitor system; Infrared camera behind the LCD; High infrared light transmittance film; Defogging algorithm; Guided filtering

1. Introduction

Fatigue and distracted driving are common causes of traffic accidents. For the sake of public transportation safety, multiple countries have begun to introduce relevant laws and regulations to promote the application of DMS. DMS will become a safety standard configuration for the new generation of intelligent vehicles.

According to the different characteristics of fatigue, there are many methods to detect, such as fatigue detection based on vehicle behavior, driver behavior, driver physiological characteristics, and multi feature fusion [1-3]. The fatigue detection based on vehicle behavior will compare the information of characteristic parameters such as vehicle speed, acceleration, vehicle yaw angle, and the degree of vehicle deviation from the lane with the standard parameters of normal driving [4]; Fatigue detection based on driver behavior will identify the driver's fatigue state based on factors such as steering wheel grip strength [5], EEG signals [4], blink frequency [6-8], etc. when the driver operates the vehicle.

At present, one of the major solutions of DMS is that placing infrared cameras on the dashboard, steering column, left and right A-pillars, or rearview mirrors of the car, and its built-in infrared light source illuminates the driver. The reflected infrared light from driver is received by the infrared camera to obtain the driver's image. By facial recognition, alignment [9-10], and calculating the relative positions of the facial landmarks, it is possible to determine whether the driver yawns with their mouth opened, sleeps with their eyes closed, distract themselves with their head turned, and so on.

2. New DMS System

In order to miniaturize the car frame structure and maximize the interior space, we integrated the infrared camera behind the dashboard LCD and used two infrared light sources to be integrated into the left and right A-pillars respectively. In order to adapt the integration of LCD and camera to any vehicle model, the infrared camera can be placed at any position behind the LCD to capture images, which requires that the pixel structure at any position on the LCD be the same. This periodic pixel structure benefits display effect, but causes the images captured by infrared camera are inevitably extremely blurry.

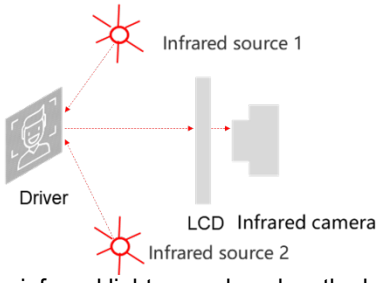

Fig. 1. Two infrared lights are placed on the left and right sides of the driver, shining on the driver's face and reflecting, passing through the LCD and being received by the infrared camera behind it.

To solve the problem of blurry images captured by the infrared camera behind the LCD, we first analyze the physical process that occurs when light passes through the LCD, and then seek targeted image processing methods based on this process. After the image is clear, general facial recognition and motion detection algorithms can be used to implement the DMS system.

LCD is composed of polarizer, liquid crystal layer, driving circuit, light guide plate, etc. For infrared light, polarizer, liquid crystal layer, light guide plate, etc. are like layers of thin films with different transmission coefficients. When infrared light passes through these thin films, it undergoes reflection, refraction, absorption, and transmission. The image becomes increasingly blurry with each layer of thin film. The TFT array is similar to a lattice grating, with a gap size close to the wavelength of infrared light. The fluctuation of light cannot be ignored, and diffraction phenomena are more obvious.

ISSN 0097-996X/25/5602-0983-$1.00 © 2025

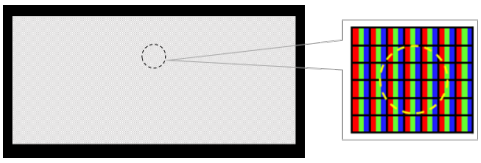

Fig. 2. The pixels at the LCD to the corresponding infrared camera are arranged in a periodic structure without any modification to the pixel arrangement.

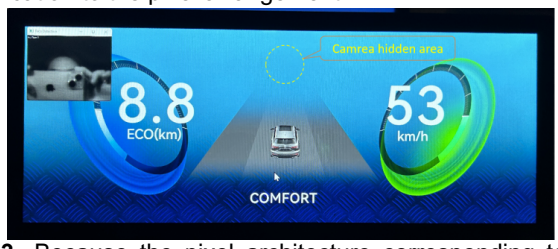

Fig. 3. Because the pixel architecture corresponding to the infrared camera area is the same as elsewhere, without sparse processing, the display uniformity is good.

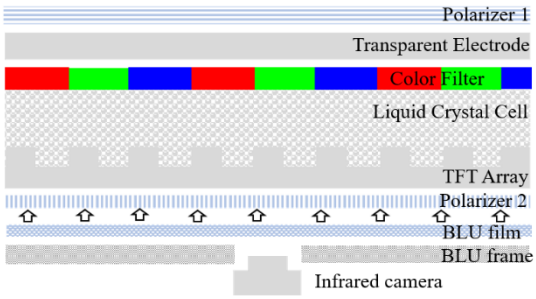

Fig. 4. The display of an LCD is the process of white light being sent from the light guide plate to polarizer 1; The infrared light infrared camera get is sent from polarizer 1 to the light guide plate. LCD need to be able to control the output of visible light, while also minimizing the loss of infrared light that can be received by infrared cameras. This requires significant differences in the processing of visible and infrared light by the materials inside the LCD [11-15].

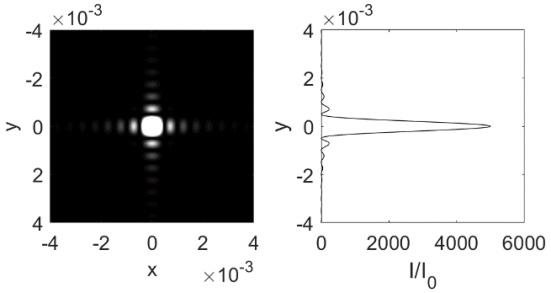

Fig. 5. The TFT array is similar to a grid grating, which is arranged by square holes. The left image shows the diffraction spot of a point light source passing through a square hole [16]; The right figure shows the intensity distribution I/I_0 of the diffraction spot in the y-direction.

The front image of the LCD passes through multiple layers of thin film and then through a grid grating. When each beam does not meet the coherence condition, each pixel in the image is blurred into a light spot similar to Figure 5. The image after the grating is the result of the superposition of numerous blurred light spots. In the grid diffraction spot, the zero order spot has the greatest impact on image quality, while the light intensity and area of other order spots are relatively small. The intensity of the zero order spot is affected by the intensity of the light source in front of the grating, and its area is related to the size of the TFT array and the wavelength of infrared light. By designing the TFT array size and light source wavelength, the diffraction spot of the image can be minimized as much as possible. This way, the outline of the image passing through the grid grating is still preserved, but the edges are blurred and bolded. This blurring is similar to being caused by a layer of fog, so in terms of performance, we can approximate a grid grating as a layer of medium with different coefficients.

3. Simplified Model

Based on the analysis of the reasons for the blurry image from the infrared camera behind the LCD, we simplified the LCD into layers of thin films with different media during image processing. The image becomes more blurred with each layer of thin film. The image processing process we recommend is to remove fog layer by layer from the image received by the infrared camera behind the screen, ultimately obtaining a clear image.

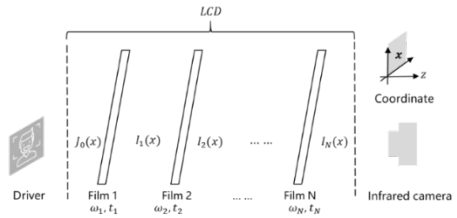

Fig. 6. The LCD is simplified into layers of thin films, and the driver's image will become more blurred after passing through each layer of film.

The reflected light from the driver is absorbed by particles in the medium during the transmission of each layer of thin film, resulting in a decrease in image brightness; Reflected light is scattered by opaque substances to form stray light, resulting in a decrease in image contrast. We will first analyze the process of light passing through a thin film, and then discuss the algorithm for removing the film.

Reflected light gradually decays due to absorption and scattering as the transmission distance increases during propagation. The light intensity is the luminous flux per unit cross-sectional area. Starting from position z, the intensity changes for each distance dz transmitted as follows

$$\frac{dE(z,\lambda)}{E(z,\lambda)} = -\beta(\lambda)dz\,. \tag{1}$$

$E(z,\lambda)$ represents the attenuated light intensity, $\beta(\lambda)$ is the attenuation coefficient. The wavelength λ of the light received by the infrared camera is within the infrared light band. The original beam intensity is denoted as $E_0(\lambda)$, and the attenuated intensity at z is

$$E(z,\lambda) = E_0(\lambda)e^{-\beta(\lambda)z}\,. \tag{2}$$

Light is not only absorbed but also scattered within the thin film. The scattered path and distance are random, and the scattered light has lost the information of the target image. While the infrared camera not only receives attenuated light with image information, but also scattered light without image information. We classify this scattered

light into ambient light and label it as

$$\mathrm{E_e}(\mathrm{z},\lambda) = E_{e0}(\lambda)\left(1 - e^{-\beta(\lambda)z}\right). \quad (3)$$

The infrared camera received a total of attenuated reflected light and ambient light including scattered light. Many light beams illuminate a plane perpendicular to the z-axis, presenting an image.

$$\mathrm{I}(x) = J(x)t(x) + A\left(1 - t(x)\right). \quad (4)$$

$\mathrm{I}(x)$ is the image after a layer of medium, $J(x)$ is the original image, transmittance $t(x) = e^{-\beta(\lambda)x}$, and ambient light $\mathrm{A} = E_{e0}(\lambda)$.

A model based on the absorption and scattering of a single thin film, considering the absorption and scattering models of multiple thin films with different media. The image after the i-th layer of film is

$$\mathrm{I_i}(x) = I_{i-1}(x)t_i(x) + A_i\left(1 - t_i(x)\right). \quad (5)$$

The image after the last layer, $\mathrm{I_N}(x)$ is received by an infrared cam era, and the shape information with the target object is $\mathrm{I_{N-1}}(x)$. The shape information with the target object in $\mathrm{I_{N-1}}(x)$ is $\mathrm{I_{N-2}}(x)$, an d so on. It is necessary to trace back layer by layer to obtain $J_0(x)$.

4. Image Processing

Firstly, consider the method of solving the original image before a thin film layer. From the equation (4), this is similar to a problem of blurring foggy images. There are many methods for defogging, such as dark channel priors [17], Retinex [18], and deep learning based defogging algorithms [19].

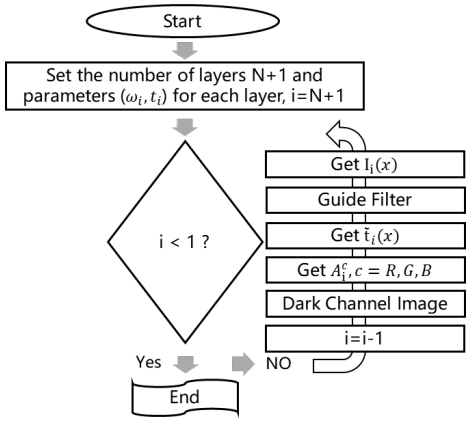

Fig. 7. Image processing algorithm flowchart.

Due to the scarcity of infrared light in the environment, the infrared light source mainly illuminates the driver's body, while other parts are dark. In this way, the infrared camera behind the LCD always captures images with some pixels having at least one color channel with a very low grayscale value, which satisfies the prior condition [17]

$$\mathrm{J^{dark}}(x) = \min_{\mathrm{c}}\left(\min_{y\in\Omega(x)}\left(J^c(y)\right)\right) = 0 \quad (6)$$

for the dark channel defogging algorithm. Therefore, we chose the dark channel defogging method to increase the clarity of the image after a thin film. In equation (4), $\Omega(x)$ is the local pixel centered around x, and J^c is a color channel of the $J(x)$ image $(\mathrm{c} = \mathrm{R}, \mathrm{G}, \mathrm{B})$. Next, we will solve equation (4), and add parameters that characterize the properties of the thin film for debugging.

Filter equation (4) by finding the minimum value on both sides and divide it by A^c on both sides

$$\min_{y\in\Omega(x)}\left(\frac{I^c(y)}{A^c}\right) = \tilde{\mathrm{t}}(\mathrm{x}) \min_{y\in\Omega(x)}\left(\frac{J^c(y)}{A^c}\right) + \left(1 - \tilde{\mathrm{t}}(x)\right). \quad (7)$$

Find the minimum value for each color channel on both sides of the equation again

$$\min_{\mathrm{c}}\left(\min_{y\in\Omega(x)}\left(\frac{I^c(y)}{A^c}\right)\right) = \tilde{\mathrm{t}}(x) \min_{c}\left(\min_{y\in\Omega(x)}\left(\frac{J^c(y)}{A^c}\right)\right) + \left(1 - \tilde{\mathrm{t}}(x)\right). \quad (8)$$

Based on the prior theory of dark channels (6), the transmittance coefficient $\tilde{\mathrm{t}}(x)$ is

$$\tilde{\mathrm{t}}(x) = 1 - \min_{\mathrm{c}}\left(\min_{y\in\Omega(x)}\left(\frac{I^c(y)}{A^c}\right)\right). \quad (9)$$

When processing each layer of film from the back of the screen to its front, there is still a film in front, so the image during the processing always retains a bit of fog. Introduce a constant parameter $\omega(0 < \omega < 1)$ to control the degree of film removal, so

$$\tilde{\mathrm{t}}(x) = 1 - \omega \min_{\mathrm{c}}\left(\min_{y\in\Omega(x)}\left(\frac{I^c(y)}{A^c}\right)\right). \quad (10)$$

Then obtain the top 0.1% brightness pixel positions from the dark channel map, and find the grayscale with the highest brightness at these pixel positions in the $\mathrm{I}(x)$ image as the A^c value. In order to protect the edges as much as possible without losing the contour of the image, it is recommended to use guided filtering to process $\tilde{\mathrm{t}}(x)$. In the calculation process, if the transmittance $\tilde{\mathrm{t}}(x)$ tends to 0, $J(x)\tilde{\mathrm{t}}(x)$ also tends to 0, thus losing the information of the original image. Therefore, a parameter $\mathrm{t_0}$ is introduced as the lower limit of transmittance. The smaller $\mathrm{t_0}$, more light passes through the film, and $\mathrm{t_0}$ is also a parameter related to the transmittance of the film. By combining the equation (4) and (10), we can obtain the image before a thin film layer

$$\mathrm{J}(x) = \frac{\mathrm{I}(x) - \mathrm{A}}{\max\left(\tilde{\mathrm{t}}(x), \mathrm{t_0}\right)} + A. \quad (11)$$

Given the image processing method for removing single-layer films, it can be inferred that the process of removing multi-layer films. The output image $\mathrm{I_i}(x)$ of the next layer of film i is used as the input image of the previous layer of film i-1, and the solution is solved forward layer by layer until the original image $J_0(x)$ is obtained.

5. Basic Function Implementation of DMS

We use the above multi-layer defogging method to make the infrared image behind the LCD clearer, and then use the cascaded residual tree method for face alignment [9], calibrating 68 feature points of the face.

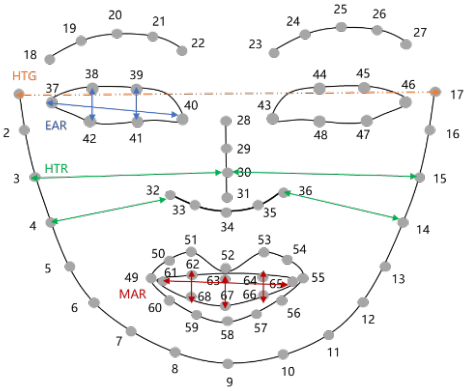

Fig. 8. 68 facial feature points. From the two-dimensional coordinates of 68 points, four states of the driver can be captured: mouth open, eyes close, head tilt, and head turn.

Determine whether the driver is fatigued and distracted by the relative position of facial landmarks. If there is a closed eye movement, it is very fatigue; If yawning occurs, it indicates fatigue; If the head turns in another direction, it is distraction. The characteristic of closed eyes is that the opening amplitude of the eyes is extremely small and the blink frequency is extremely low. Tereza Soukupová suggested using eye aspect ratio (EAR) parameters to characterize the amplitude of eye opening [6].

$$\mathrm{EAR} = \frac{||\mathrm{P}_{38} - \mathrm{P}_{42}|| + ||\mathrm{P}_{39} - \mathrm{P}_{41}||}{2||\mathrm{P}_{37} - \mathrm{P}_{40}||}. \quad (12)$$

$$||\mathrm{P_i} - \mathrm{P_j}|| = \left(\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2\right)^{\frac{1}{2}}, \quad (13)$$

$||\mathrm{P_i} - \mathrm{P_j}||$means the distance of point i and j. The characteristic of yawning is that the mouth is particularly wide open, exceeding the usual range of speech. Drawing inspiration from Tereza Soukupová's quantification of the degree of eye opening, the degree of mouth opening is quantified using the mouth aspect ratio (MAR) parameter [6].

$$\mathrm{MAR} = \frac{||\mathrm{P}_{62} - \mathrm{P}_{68}|| + ||\mathrm{P}_{63} - \mathrm{P}_{67}|| + ||\mathrm{P}_{64} - \mathrm{P}_{66}||}{3||\mathrm{P}_{61} - \mathrm{P}_{65}||}. \quad (14)$$

When turning the head, the distance between the edges of the left and right faces and the nose is different. The head turning ratio

$$\mathrm{HTR} = \left|\frac{D_{\mathrm{left}} - D_{\mathrm{right}}}{D_{\mathrm{left}} + D_{\mathrm{right}}}\right|, \quad (15)$$

$$D_{\mathrm{left}} = ||\mathrm{P}_3 - \mathrm{P}_{30}|| + ||\mathrm{P}_4 - \mathrm{P}_{32}||,$$
$$D_{\mathrm{right}} = ||\mathrm{P}_{14} - \mathrm{P}_{36}|| + ||\mathrm{P}_{15} - \mathrm{P}_{30}||.$$

When tilting the head, the gradient of the line connecting point 1 and point 17 is relatively large. The head tilting gradient

$$\mathrm{HTG} = \left|\frac{\mathrm{x}_{17} - \mathrm{x}_1}{\mathrm{y}_{17} - \mathrm{y}_1}\right|. \quad (16)$$

$(\mathrm{x_i}, \mathrm{y_i})$ is the coordinate point of i in the image.

Fig. 9. When turning the head, there is a significant difference in the distance between the left and right facial edges and the nose. When not turning, the difference is relatively small.

The abnormal state of the head includes turning, tilting, lifting, and lowering, which requires three coordinate axes x, y, and z to represent. Tilting the head refers to the rotation of the head about the x-axis, turning the head refers to the rotation of the head about the z-axis, lifting and lowering the head refers to the rotation of the head about the y-axis. But currently we have not found a method that can effectively distinguish head up and head down on two-dimensional faces, which is a drawback of this solution.

The setting of thresholds EAR, MAR, HTR, and HTG requires debugging. They vary depending on the image quality, and greatly affect the recognition ratio of opening mouth, closing eyes, turning head, and tilting head.

6. Experimental Results

We strive to improve the image quality of the infrared camera behind the LCD to facilitate the implementation of DMS. The evaluation of general image quality improvement methods is mainly based on subjective feelings. One of the requirements of DMS is a high facial recognition ratio. The multi-layer defogging image processing method serves the implementation of DMS, and its objective value will be reflected in the improvement of facial recognition ratio. Therefore, we propose using face recognition ratio (FRR) to evaluate the performance of multi-layer defogging image processing methods.

Firstly, place cameras in front and behind the LCD to record videos of the driver driving normally multiple times in bright and dark environments, driving fatigue (yawning and naked eyes), and driving distraction (turning and tilting the head). Then have the driver put on glasses and record the same action again. Subsequently, the infrared camera behind the LCD underwent multi-layer defogging treatment. Finally, compare the FRR in these situations. FRR is the ratio of the number of frames in which a face is detected to the number of frames that actually include the face.

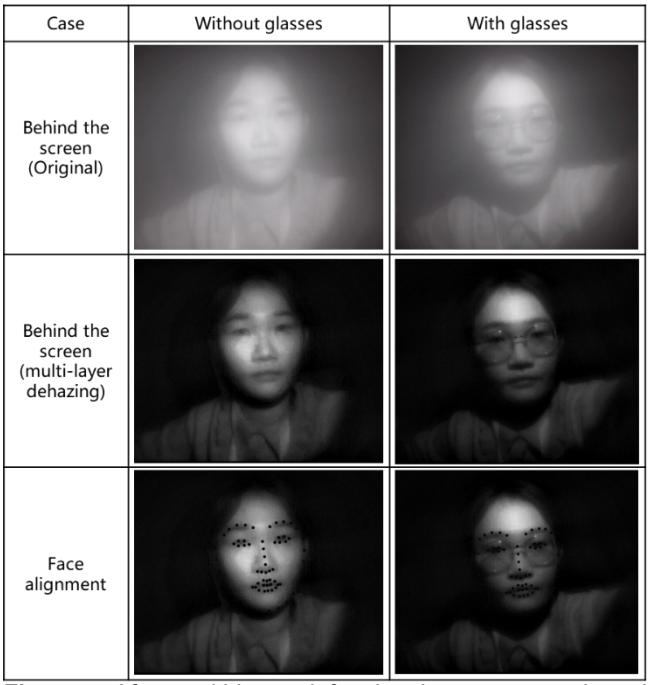

Fig. 10. After multi-layer defogging image processing, the image becomes clearer and the face alignment is more accurate.

Generally, infrared cameras do not go through an LCD, resulting in very clear image quality and no need for image processing. The infrared camera placed behind the LCD needs to undergo image processing. The multi-layer defogging algorithm improves the FRR by nearly 60%, making the FRR with LCD extremely similar to that without LCD. However, the FFR when wearing glasses lags behind that without an LCD by nearly 8%. This is because when the head is turned to the extreme left or right position, a large part of the infrared light source is reflected by the glasses, and after passing through the LCD, a large diffraction spot is generated, which obscures most of the face and leads to facial recognition failure. This indicates that the detection range of the new DMS is relatively limited. In order to make the image received by the infrared camera behind the screen appear to

have no LCD, our next step is to solve the serious spot problem caused by strong light diffraction.

From Figure 10, it can be seen that the image after the multi-layer defogging algorithm is perceived to be clearer. When wearing glasses, the infrared light source will always be reflected, but when the angle of the infrared light source is appropriate, the strongly reflected light will not be received by the infrared camera, so there will be no serious light spot and it can be detected normally.

7. Conclusion

The proposed under-screen camera solution marks a new milestone to enhance screen compatibility and reduce costs for terminal factory. The camera is well hidden below the screen, and the corresponding screen area is visible, indistinguishable from other display areas. This eliminates the display defects associated with previous blind hole solutions. On the precondition that the infrared transmittance of the monitor is 30%, the multi-layer defogging algorithm can effectively make the infrared image clearer and improve the FRR by about 60%. However, due to the limitation of strong light diffraction, the detection range of the new DMS for drivers is relatively limited. In the future, we will strive to design suitable light paths and modify pixel arrangements to avoid serious spot problems caused by strong light diffraction.

References

[1] MA Cuaya, JF Osorio-Valencia and RE Vasquez, "A survey of fatigue detection methods: State-of-the-art and challenges", *Computers in Human Behavior* (2020).

[2] Qin, Y., Lyu, H. & Zhu, K. "Driver fatigue detection method based on multi-feature empirical fusion model". *Multimed Tools Appl* (2024).

[3] Tasawor Ahmed Sofi, Shabana Mehfuz, "Drowsiness and fatigue detection using multi-feature fusion", *2023 International Conference on Recent Advances in Electrical, Electronics & Digital Healthcare Technologies (REEDCON),* pp.672-675 (2023).

[4] Zhendong Lan, Jian Zhao, Pengbo Liu, Chi Zhang, Nana Lyu, Lie Guo. "Driving fatigue detection based on fusion of EEG and vehicle motion information", *Biomedical Signal Processing and Control*, vol. 92 (2024).

[5] Rui Li, Yingjie Victor Chen, Linghao Zhang, "A method for fatigue detection based on Driver's steering wheel grip", *International Journal of Industrial Ergonomics,* vol.82 (2021).

[6] Tereza Soukupová, "Eye-Blink Detection Using Facial Landmarks", *Czech Technical University in Prague* (2016).

[7] Akihiro Kuwahara, Kazu Nishikawa, Rin Hirakawa, Hideaki Kawano, Yoshihisa Nakatoh, "Eye fatigue estimation using blink detection based on Eye Aspect Ratio Mapping(EARM) ", *Cognitive Robotics*, vol.2(2022).

[8] A. Kuwahara, R. Hirakawa, H. Kawano, K. Nakashi and Y. Nakatoh, "Eye Fatigue Prediction System Using Blink Detection Based on Eye Image", *2021 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 1-3 (2021)

[9] Lai, Hanjian & Xiao, Shengtao & Cui, Zhen & Pan, Yan & Xu, Chunyan & Yan, Shuicheng. " Deep Cascaded Regression for Face Alignment", 10.48550/arXiv.1510.09083 (2015).

[10] Shi, Baoguang and Bai, Xiang and Liu, Wenyu and Wang, Jingdong, " Deep Regression for Face Alignment", *IEEE Transactions on Neural Networks and Learning Systems* (2014).

[11] WangZ ,ChiL ,ChangY , et al . " Diffraction Issues of Under Display IR Sensor in AMOLED Displays".*SID Symposium Digest of Technical Papers,* 55(1):738-741(2024).

[12] Yang M ,Gao K Z ,Xu X C , et al. "Design and fabricate of Under‐Display Camera OLED panel".*SID Symposium Digest of Technical Papers*, 54(1):337-340 (2023).

[13] Hengcheng Z ,Wei X , Shuai M , et al. "Ne-liquefaction system development for the 940 nm infrared optical characteristics measurement". *Cryogenics*, (prepublish): 103411 (2021).

[14] Lee J ,Kim K . "An Experimental Study on the Penetration of 850nm and 940nm Infrared Radiation into Porcine Tissues". *Indian Journal of Public Health Research & Development,* 10(5): 1086-1089(2019).

[15] Zhenhua Z .14.4: "Diffraction Simulation of Camera Under Display".SID Symposium Digest of Technical Papers ,5293-96(2021).

[16] Liu Dong-zhou, "Numerical simulation of square hole Fraunhofer diffraction", *College Physics* (2011).

[17] Kaiming He, Jian Sun and Xiaoou Tang, "Single image haze removal using dark channel prior," *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1956-1963 (2009).

[18] J. Wang, K. Lu, J. Xue, N. He and L. Shao, "Single Image Dehazing Based on the Physical Model and MSRCR Algorithm", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 9, pp. 2190-2199(2018).

[19] Ravi Raj Choudhary, K K Jisnu, Gaurav Meena, "Image Dehazing Using Deep Learning Techniques", *Procedia Computer Science,* vol.167, pp.1110-1119 (2020).