AI-Based Frontalization of Low-Angle Faces for Questioner-Overlay Presentation Systems

Akira Shiraki
Nagoya Institute of Technology, Japan
a.shiraki.048@stn.nitech.ac.jp

Kenji Funahashi
Nagoya Institute of Technology, Japan
kenji@nitech.ac.jp

***Abstract*—During presentations, questions may come from individuals who are not easily identifiable. In large venues, this makes it difficult for presenters and audiences to visually identify the questioner. To address this issue, a presentation support system has been proposed that overlays an image of the questioner using an omnidirectional camera integrated into a microphone. However, because the camera is positioned below, it captures faces from a low angle, which creates a sense of visual inconsistency when displayed as-is. This paper proposes an AI-based method for frontalizing low-angle face images and applies it to the presentation support system. The proposed method is expected to improve both interest in and understanding of the questioner.**

***Index Terms*—face frontalization, generative AI, presentation support system, omnidirectional camera, head pose estimation**

I. Introduction

In recent years, the number of opportunities for presentations using projectors has increased, and interest in presentation support systems has grown [1]. In large venues, questions are typically asked using handheld microphones, making it difficult for presenters and audiences to visually identify the questioner. To address this issue, a presentation support system has been proposed that integrates an omnidirectional camera into the microphone to continuously capture the questioner regardless of camera position, and overlays the questioner's face onto the screen, thereby enhancing the audience understanding and engagement [2] (Fig. 1). However, because the camera is located below the face, it captures facial images from a low angle, resulting in an unnatural appearance when displayed as-is. This unnatural appearance can reduce the audience's ability to intuitively recognize the questioner and may negatively affect engagement during the presentation. Previous studies have attempted to frontalize such low-angle face images using 3D shape models [3]. However, because the model does not perfectly match the individual's face, the quality of the frontalized images remains limited. Recent advances in deep learning have enabled high-quality face generation and transformation. This study proposes an AI-based method for frontalizing low-angle face images using deep learning-based image transformation techniques, aiming to achieve higher-quality frontalization in the presentation support system.

II. Implementation and System

Face frontalization is performed using FacePoke [4], a face image transformation tool based on generative AI. FacePoke enables the manipulation of facial orientation and expression, allowing users to generate face images with desired configurations through interactive operations (Fig. 2). The proposed method performs automatic frontalization in three stages. First, facial landmarks are detected from low-angle face images captured by the camera using the Dlib C++ library. Next, the head pose relative to the camera is estimated using the detected landmarks and the Perspective-n-Point (PnP) pose computation algorithm in OpenCV. Finally, based on the estimated pose, FacePoke is applied to transform the low-angle face into a frontal view. The novelty of the proposed method lies in the automatic estimation of head pose and its integration with generative face transformation to enable real-time frontalization in a presentation setting.

Fig. 1. An omnidirectional camera integrated into a microphone

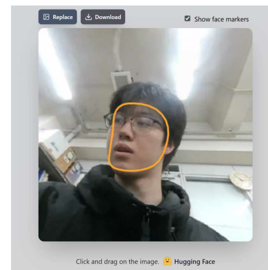

Fig. 2. FacePoke interface for manipulating facial orientation and expression

[Face Detection and Landmark Extraction] Facial landmarks are extracted from captured low-angle face images using the Dlib library. The detected landmarks include key facial points such as the eye corners, nose tip, and jawline, which are used as reference points for head pose estimation.

[Head Pose Estimation] The PnP algorithm estimates the 3D head pose relative to the camera by mapping the detected 2D facial landmarks onto a predefined 3D face model. This process yields three rotation angles: pitch (vertical rotation), yaw (horizontal rotation), and roll (in-plane rotation), which represent the deviation from a frontal face orientation.

[Automatic Face Frontal Correction] Based on the estimated pose angles, FacePoke is applied with parameters specifying a frontal target pose. The generative model transforms the low-angle face image into a frontal view while preserving the subject's identity and facial expression. The frontalized image is then integrated into the presentation support system for display.

III. Evaluation

We conducted two evaluations to assess the effectiveness of the proposed method.

[Evaluation 1] First, we compared images corrected using the conventional method and the proposed method to evaluate their effectiveness. For this evaluation, low-angle face images with pitch angles of 30°, $45°$, and $60°$ were used. The comparison results are shown in Fig. 3. Each figure presents, from left to right, the original low-angle image, the result of the conventional method, and the result of the proposed method. As shown in Fig. 3 (top), for a pitch angle of 30°, the conventional method produces visible artifacts around the boundary between the face and the background, whereas the proposed method yields a more natural appearance. In Fig. 3 (middle: $45°$ and bottom: $60°$), this issue becomes more pronounced in the conventional method, with increasing distortion and unnatural boundaries. In contrast, the proposed method maintains a relatively natural boundary between the facial region and the background even as the pitch angle increases. These results indicate that the proposed method achieves more natural face frontalization than the conventional method, particularly for larger pitch angles. Although the evaluation is qualitative, consistent improvements were observed across all tested pitch angles.

[Evaluation 2] Next, we evaluated the effectiveness of the proposed method within the presentation support system. Evaluators conducted mock presentations and Q&A sessions to compare the impressions of AI-frontalized images and uncorrected images. Fig. 4 shows an example of the evaluation setup and display. The left image shows a mock presentation in a classroom environment, while the right image shows the presentation screen with the frontalized questioner overlaid. The use of frontalized images appeared to increase audience interest in the Q&A session and improve understanding of the questioner. However, due to a decrease in frame rate caused by the frontalization process, the displayed video appeared less smooth, making it difficult to perceive subtle facial expressions. Furthermore, the effectiveness of the proposed method was not strongly reflected in the presenters' evaluations. This is likely because the evaluation environment allowed presenters to visually identify questioners directly, reducing the impact of the overlaid frontalized images. The evaluation was conducted with a limited number of participants, and further quantitative studies are required for more rigorous validation.

Fig. 3. Images with pitch angles of 30°, $45°$, and $60°$ (top to bottom, from left: original, frontalized by the conventional method, frontalized by the proposed method)

Fig. 4. Mock presentation and Q&A session (left) and presentation screen with the frontalized questioner overlaid (right)

IV. Conclusion and Future Work

In this study, we proposed an AI-based method for frontalizing low-angle face images captured by a microphone-integrated camera in a presentation support system. The proposed method achieves higher-quality frontalization compared to conventional approaches based on 3D models. The results of the evaluation indicate that, although the effectiveness of the proposed method was not strongly perceived by presenters, it contributed to increased audience interest and improved understanding of the questioner. This approach demonstrates the potential of combining generative AI with presentation support systems to enhance interactive communication. Future work includes improving real-time performance to address frame rate reduction and further enhancing the naturalness and familiarity of the frontalized images to increase the overall effectiveness of the system.

Acknowledgment This work was supported in part by JSPS KAKENHI Grant Number JP23K11182.

References

[1] R. Inoue, S. Shiramatsu, T. Ozono, T. Shintani, "An Interactive Presentation System Based on Feedback to a Presentation Material of an Ongoing Presentation (in Japanese)," IPSJ Journal, vol. 56, no. 10, pp. 2011-2021, Oct. 2015.

[2] Y. Kobayashi, K. Funahashi, "Superimposing Questioner on Presentation Screen Using Microphone with Whole-Sky Camera," ICAT-EGVE 2016, pp.3-4, 2016.

[3] D. Shibata, Y. Kobayashi, K. Funahashi, "Presentation Support System Overlaying the Questioner Using an Omnidirectional Camera-Integrated Microphone," NICOGRAPH2019, 2019.

[4] FacePoke, https://facepoke.ai/.