

# “It Is Hard to Remove from My Eye”: Design Makeup Residue Visualization System for Chinese Traditional Opera (Xiqu) Performers

Zeyu Xiong  
Computational Media and Arts Thrust  
The Hong Kong University of Science  
and Technology (Guangzhou)  
Guangzhou, China  
zxiong666@connect.hkust-gz.edu.cn

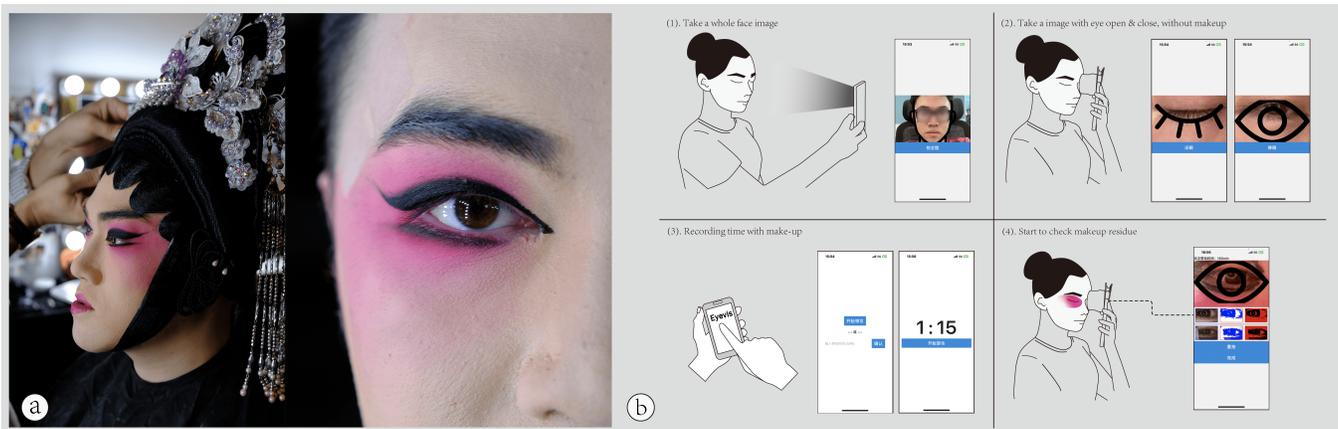
Shihan Fu  
Computational Media and Arts Thrust  
The Hong Kong University of Science  
and Technology (Guangzhou)  
Guangzhou, China  
sfu663@connect.hkust-gz.edu.cn

Yanying Zhu  
Data Science and Analysis Thrust  
The Hong Kong University of Science  
and Technology (Guangzhou)  
Guangzhou, China  
yzhu367@connect.hkust-gz.edu.cn

Chenqing Zhu  
Internet of Things Thrust  
The Hong Kong University of Science  
and Technology (Guangzhou)  
Guangzhou, China  
czhu032@connect.hkust-gz.edu.cn

Xiaojuan Ma  
Department of Computer Science and  
Engineering  
The Hong Kong University of Science  
and Technology  
Hong Kong SAR, China  
mxj@cse.ust.hk

Mingming Fan\*  
Computational Media and Arts Thrust  
The Hong Kong University of Science  
and Technology (Guangzhou)  
Guangzhou, China  
Division of Integrative Systems and  
Design & Department of Computer  
Science and Engineering  
The Hong Kong University of Science  
and Technology  
Hong Kong SAR, China  
mingmingfan@ust.hk



**Figure 1: Overview of Makeup Residue Visualization System (The photo is used with consent). (a) Chinese Traditional Opera Makeup, (1) - (4): Sample Usage of the system, step (1): Take an image of the whole face, step (2): Take images of eyes (open & close), step (3): Recording Makeup Wearing Time, and step (4): Eye Makeup Visualization.**

\*Corresponding Author

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

CHI '24, May 11–16, 2024, Honolulu, HI, USA

## ABSTRACT

Chinese traditional opera (Xiqu) performers often experience skin problems due to the long-term use of heavy-metal-laden face paints. To explore the current skincare challenges encountered by Xiqu performers, we conducted an online survey (N=136) and semi-structured interviews (N=15) as a formative study. We found that

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 979-8-4007-0330-0/24/05  
<https://doi.org/10.1145/3613904.3642261>

incomplete makeup removal is the leading cause of human-induced skin problems, especially the difficulty in removing eye makeup. Therefore, we proposed *EyeVis*, a prototype that can visualize the residual eye makeup and record the time make-up was worn by Xiqu performers. We conducted a 7-day deployment study (N=12) to evaluate *EyeVis*. Results indicate that *EyeVis* helps to increase Xiqu performers' awareness about removing makeup, as well as boosting their confidence and security in skincare. Overall, this work also provides implications for studying the work of people who wear makeup on a daily basis, and helps to promote and preserve the intangible cultural heritage of practitioners.

## CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; **Assistive technologies**.

## KEYWORDS

makeup, Chinese traditional opera, computer vision, mobile computing, intangible cultural heritage, interactive design

### ACM Reference Format:

Zeyu Xiong, Shihan Fu, Yanying Zhu, Chenqing Zhu, Xiaojuan Ma, and Mingming Fan. 2024. "It Is Hard to Remove from My Eye": Design Makeup Residue Visualization System for Chinese Traditional Opera (Xiqu) Performers. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 16 pages. <https://doi.org/10.1145/3613904.3642261>

## 1 INTRODUCTION

Stage performers have a high risk of suffering from occupational health problems [30]. Especially for the performing arts which require high-quality visual effects, actors might need to wear heavy makeup frequently to achieve such effects and are therefore heavily exposed to cosmetics [25]. Wearing makeup for a long period of time as well as inappropriate and incomplete makeup removal can easily cause skin irritation and a host of other problems, especially if the makeup contains heavy metals [74]. One of the representative user groups is Chinese traditional opera (Xiqu) performers. Xiqu is one type of the most significant intangible cultural heritage (ICH) under the United Nations Educational, Scientific and Cultural Organization (UNESCO) list [71]. As intangible cultural practitioners, Xiqu performers have made great contributions to society. According to the investigation from the Ministry of Culture (PRC) [57] by 2018, there are 348 genres of Xiqu and over 10,278 troupes registered, and there are over 300,000 Xiqu performers performing actively on stage, based on the survey of China Theater Association [4] by 2019.

Xiqu is also a form of musical stage art that highly demands visual effects [18, 43, 81]. Specifically, thick face paints are usually used with various colors in terms of different roles of play (e.g. Sheng, Dan, Jing, Mo, Chou [61]) to achieve considerable visual effects [22] and expression of emotions [91]. However, face paint consists of various types of heavy metals [5, 10, 16], and the long wear of face paint is extremely harmful to the skin [74]. It is usually used by Xiqu performers for about a quarter of the year on average [75]. Aside from the necessity of the performance, wearing make-up for a long time can also be due to incomplete makeup

removal, which might lead to skin problems such as pore clogging, increased oil secretion, skin allergies, and acne vulgaris [38]. Therefore, the skin health of Xiqu performers caused by face paint makeup needs more attention. To reduce the hassle of applying and removing traditional makeup, researchers have proposed new materials for makeup [35, 56] and interactive makeup methods [67]. However, little is known about preventing skin damage during the makeup removal process from the user's perspective. To this end, we propose a research question (RQ1): **What are the practices and challenges of making-up removal among Xiqu performers?**

To answer RQ1, we conducted a formative study in which we first performed a survey that understands the fundamental data of Xiqu performers' makeup removal process. Among 136 responses, we selected 15 representative participants to conduct a semi-structured interview to understand the detailed challenges and practices. We open-coded the results and found that incomplete makeup removal is the main reason, and the eye makeup is the hardest part to remove, and we derived six design considerations (DC). Based on the DCs, we further designed and implemented *EyeVis*, an interactive system that is tailored for Xiqu performers to visualize the eye makeup residue during the removal process, and track the trend of makeup-wearing time, consisting of a mobile app with a fill light & a camera lens magnifier & eye shields. To our knowledge, *EyeVis* is the first ever tool designed for eye makeup residue visualization. We then used it to investigate the effectiveness and user experience of these features by answering the second research question (RQ2): **What features and functionalities would Xiqu performers find valuable in a system designed to aid in eye makeup removal?** To answer RQ2, we conducted a 7-day deployment study where 12 Xiqu performers participated simultaneously. During the deployment study, participants were asked to complete makeup removal tasks every day and participated in the final day interview. We open-coded the data and derived our findings by data analysis. The results showed the usability of *EyeVis*, including the visualization methods to improve the efficiency of the makeup removal process from multiple perspectives and increase the awareness of the Xiqu performers. Finally, we discussed potential generalization directions for our work in terms of wider visualization area of skin, wider audiences, and wider usage for skin data collection.

In summary, we made the following contributions:

- We investigated common practices and challenges for the making-up removal process of Xiqu performers through a formative study including an online survey and semi-structured interviews, and we proposed six informative design considerations for the future design of similar skincare systems.
- Informed by the formative study, we designed and implemented an interactive prototype that allows Xiqu performers to see the makeup residue clearly through the visualization system, and see the makeup wearing time by a weekly trend.
- We conducted a 7-day deployment study with a community of Xiqu performers to understand how they used and perceived these features and presented design implications. To our knowledge, this is the first prototype that has been designed and evaluated for makeup residue visualization.

## 2 RELATED WORK

### 2.1 Xiqu Performers Suffer from Skin Problems

Standard face paint products for professionals are subjected to frequent random market controls in Western countries, in accordance with European Union (EU) cosmetic product regulations [36]. However, face paint, which is often used by Chinese traditional opera (Xiqu) performers, tends to be unregulated in China [51]. As such, their chemical composition and the degree of protection they provide to the face are questionable. As a result of wearing face paint for years and years, the skin condition of Xiqu performers is generally poor. Wang et al. [80] found that adverse reactions to face paint were present in 67.7 % of the performers, with more heavy reactions in the Dan roles (36.7 % of those with reactions) by investigating 127 Xiqu performers from Beijing Opera Theater. There were other manifestations of skin damage, such as seborrheic dermatitis, in 73 % of the performers. Wang et al. [75] conducted a questionnaire survey in which 259 Xiqu performers participated, and 53% of them reported skin diseases like skin erythema and dermatitis because of the long-term use of face paint. Besides, The condition of Xiqu's skin may get worse over the years of the performer's career. For example, Cao et al. [89] reported 2 cases of skin melanoma among Xiqu performers at the Beijing Opera Theater in 2016, and these performers both have over 30 years of experience in the field. Based on the reports and surveys mentioned above, we can find that stage performers, especially Xiqu performers do have potentially serious skin problems. Despite many studies have focused on how to lessen the toxicity of cosmetics from a chemical and medical perspective [75–77], there is currently little attention paid to the skincare of stage performers, especially Xiqu performers, in the HCI community.

To reduce the hassle of applying and removing traditional makeup, some researchers are working on new materials and interactive makeup. For example, Kao et al. [35] introduced a prototype of an interactive eyeshadow tattoo consisting of thermochromic pigments that can be activated by electronic or ambient temperature conditions. Treepong et al. [67] developed an augmented reality system that scans the human face and projects the makeup effects on the face to enhance makeup creativity. Nishimura and Siio [56] proposed iMake, an eye makeup design and printing system that utilizes ink to attach to eyes. Another alternative to traditional makeup is using masks, which are widely used in Western opera. However, Tseng and Lin [69] have investigated the difference between Xiqu makeup and masks used in Western opera, although painted faces and masks share some common functions, face paint represents the essence of Xiqu and is dissonant to be replaced by masks. Therefore, due to the uniqueness of Xiqu makeup and its irreplaceability, none of these existing works could be used in a Xiqu performer's practice, and we have yet to find a way to achieve the same visual effect without the use of face paint.

### 2.2 From Human's Perspective: Incomplete Makeup Removal

With the known risks associated with the chemicals in face paint, it is still essential to wear face paint to achieve considerable visual effect and express inherited emotion on stage [77]. As most skin

diseases are regarded as less severe or life-threatening than other organ problems, the importance of makeup removal is often overlooked. However, as skin problems are frequently visible to the public, they can have a significant impact on a patient's emotional condition [49], and even affect Xiqu performers' careers, which may result in them having to stop engaging in opera performance [25]. Aside from the fact that face paint must be worn for the show, an important human-controllable factor is how thoroughly the makeup is removed [38]. Due to the ever-evolving chemistry of cosmetics, many of them are waterproof and difficult to remove to prevent blooming [83]. If the makeup is not removed completely, it adheres to the surface of the skin and clogs the pores, causing a range of skin problems such as acne, blemishes, inflammation, etc [59]. Especially for Xiqu makeup, heavy makeup is applied to the concave and convex areas of the face (e.g. eye sockets, lips, bridge of the nose), making it more difficult to remove [39]. Although incomplete makeup removal is a crucial factor in influencing skin conditions and can be controlled by Xiqu performers, the extent to which incomplete makeup removal actually causes skin problems for Xiqu performers has not been validated by systematic research.

### 2.3 Computer-Aided Skincare Tools in HCI community

With all the challenges Xiqu performers suffer from wearing face paint, skincare has also become an important part of the post-show experience. Skincare has a strong connection with computer-aided tools [3, 73] in the intersection of medicine and HCI communities, researchers have made contributions to (1) cosmetic product development [62, 65, 90], (2) cosmetics usage and makeup tutorial [14, 41, 68, 84], (3) skin physical assessment [13, 33, 88], (4) skin condition diagnosis [27, 45, 82, 87], (5) treatment & skincare products recommendation [1, 32, 55, 66], (6) treatment outcome prediction [6, 53, 63], etc. The above work mainly involves skin imaging technology, which makes contributions to skin diagnosis and treatment by using artificial intelligence to understand various types of image information. However, we found that the focus of these existing works is mostly on the diagnosis and treatment, and how to explore the prevention of skin problems from a human-centered perspective has not yet been researched deeply.

Computer-aided techniques are also being utilized in the HCI community for everyday healthcare in the Chinese context. For example, Ding et al. [20] developed a mobile app that analyzes the face and tongue images for traditional Chinese medicine diagnosis. Wang et al. [78] proposed an AI-powered clinical decision support system in rural areas of China. From our understanding, although the research methodologies among the above works in the Chinese context have reference value, there is no such work in HCI that focuses on skincare for Xiqu performers. Although existing commercial products (e.g. facial cleansers, beauty devices) could help Xiqu performers in skincare, there is still a lack of objective metrics (e.g. augmented makeup residue visualization, skin condition metrics) that can be justified by themselves in their skincare process, various metrics still need industry experts to identify. Therefore, in addition to understanding the problems and challenges encountered by Xiqu performers in the make-up removal process, how to

use computer-aided tools to help them directly is also the focus of this work.

### 3 FORMATIVE STUDY

To answer RQ1 about understanding the practices and challenges of making-up removal mentioned in section 1, we conducted a formative study including two parts. First, we collected 136 valid responses via an online survey to investigate the makeup removal process among Xiqu performers. The primary objective of this survey was to gain initial insights into the challenges encountered during the makeup removal process. Second, we conducted semi-structured interviews (N=15) to further explore the Xiqu performers' specific needs. The findings from the formative study showcased the user need for the design of the prototype, which enables the visualization of makeup removal residue during makeup removal and obtains feedback on makeup duration and wearing trends.

#### 3.1 Online Survey

To gain deeper insights into the makeup removal process among Xiqu performers, we first designed a comprehensive online survey and distributed it through Xiaohongshu<sup>1</sup>, which is currently one of the most popular social media platforms in China. We posted posters on the platform with specific tags (i.e. #Xiqu, #Chinese Opera, #Makeup, #Makeup Removal) for the online platform recommendation to targeted users, and we also sent direct messages to accounts that identified as Xiqu performers. We collected 178 responses and screened 136 valid results by checking the reasonable completion time, the consistency of answers to repetitive questions, and setting some questions regarding professionalism (e.g. choice of make-up tools, order of applying face makeup, mechanism of make-up touch-up after long performances, etc.). The online survey covered various aspects related to makeup removal, including participants' demographic information (gender, age, years of performing experience, and type of play performed), frequency and duration of wearing face painting, perception of painting's impact on the skin, ease of removing makeup from different facial areas, and skin problems encountered.

#### 3.2 Semi-Structured Interviews

To validate the findings from the online survey and identify other potential user needs, we conducted online interviews with 15 Xiqu performers (5 males and 10 females), with varying years of performing experience (for specific demographic data of participants, see Appendix A). We recruited participants by contacting participants in the online survey who were interested in participating in our follow-up interviews. Some of the participants who contacted us through the information on our posters also participated in the interviews. Before the interview, every participant signed a consent form about collecting data for research purposes, and the consent forms passed the university's ethical review.

The interviews were semi-structured, lasting between 40 and 60 minutes. The interview questions focused on the process participants experienced to remove their makeup, the challenges they encountered, and how they believed face paint affected their skin

conditions. The interviews were audio-recorded with the participant's consent, and each participant received a compensation of RMB 100 at the end of the interview.

#### 3.3 Data Analysis

Our data consisted of online survey data as well as audio recordings from semi-structured interviews. The audio recordings were transcribed into a text script. To analyze the interview data, we followed the thematic analysis [58] method. Four researchers first read through the script to reach an overall understanding, then the open-coding procedure was carried out independently by the two co-authors once they had familiarized themselves with the data. Then, following a weekly discussion, all co-authors in the research team shared individual coding results, discussed interpretations of the data, and achieved consensus on the final coding result.

#### 3.4 Findings

##### 3.4.1 Makeup Removal Process.

*Challenges in Eye Makeup Removal.* Many Xiqu performers have expressed that removing eye makeup is the most challenging part of the makeup removal process (N=14). First, in opera performances, emphasis is placed on facial features, and multiple layers of oil-based face paints are applied to the eyes to enhance their prominence. Second, due to the concave structure and presence of wrinkles around the eyes, repeated friction is required to remove the oil-based face paints, which results in a longer time spent on the makeup removal process. (P5: "The eyeliner in Xiqu requires the oil-based face paints to be applied inside the eyelashes, very close to the eyes, which makes it difficult to clean them completely"). Besides, the skin in the eye region is more delicate compared to other regions, and there is a risk of getting chemicals in the eyes. Therefore, it is crucial to prioritize providing assistance for removing eye makeup specifically for performers' needs - **DC1: The prototype should prioritize the eye area to aid in the removal process.**

*Invisibility of Makeup Residue to Naked Eyes.* When asked about how Xiqu performers determine whether they have completely removed their makeup, all participants (N=15) mentioned relying on a subjective indicator. They either visually inspected the mirror for any remaining makeup with their naked eyes or checked if the towel used to wipe their eyes still had any residual color. (P8: "After removing my makeup, I would use a facial towel to wipe my face. If there is no color on the towel, it means I have cleaned it thoroughly"). While these methods can be effective, a portion of the performers (N=10) expressed concerns about not being entirely confident in the thoroughness of their makeup removal. (P11: "Sometimes I have to wash my face again with facial cleanser multiple times after returning home"). The standard for judging residual makeup should provide intuitive and objective references. This derived **DC2: The prototype should provide an objective and intuitive representation, allowing for easy comparison between the makeup and no-makeup conditions.**

*Difficulty of Removing Black & Pink face paints.* A unique finding is that despite the different types of operas and roles in China, most performers (N=13) still use pink and black oil-based face paints for their eye makeup. Compared to the flesh-colored face

<sup>1</sup><https://www.xiaohongshu.com/>

paints used for the base makeup, they are more concerned about removing the pink and black face paints. (P3: "The oil-based face paints on the face can be easily wiped off, but the black and pink face paints tend to smudge together because of their dark colors"). Additionally, due to the similarity between black face paints and the color of eyelashes and the difficulty of keeping the eyes open during makeup removal, performers have to observe the black face paints on their eyes multiple times. (P4: "Makeup removers can easily get into my eyes, and it's challenging to see the remaining black face paints clearly"). Therefore, the visual representation of residual makeup should offer multi-modal color visualization options - **DC3: The prototype should be multi-modal to augment black and pink face paints.**

### 3.4.2 Makeup Removal Condition.

*Constraints in Space and Time.* According to most participants reported (N=12), performers usually remove their makeup at the theater. Since performers often go to various venues and not necessarily in a fixed theater, performers need to bring their own makeup and removal tools to different locations, which require portable tools. Additionally, the makeup removal space provided by theaters is sometimes limited, then performers need to queue for their turn to use the designated area. Consequently, they have a restricted amount of time for makeup removal after the performance ends. (P11: "Some theaters have insufficient makeup removal facilities, therefore, we need to queue up and do it quickly"). Therefore, the design of the prototype should take into consideration the practical constraints of the makeup removal space and time available - **DC4: The prototype should be portable, low-cost, and user-friendly to accommodate the limited space and time available for makeup removal.**

*Limited Lighting Conditions in Theatres.* The backstage conditions at the theater are often unstable, as some performers mentioned that they frequently lack adequate conditions to remove their makeup due to poor lighting in the backstage area. More specifically, the lighting is dim, making it difficult for them to see the actual situation with the naked eye (N=9), (P15: "Some theater backstage areas have poor conditions with dim lights, making it hard to see subtle remnants of makeup on the face"). Therefore, the design of the prototype should provide a stable external lighting environment for the performers, free from the constraints of the actual lighting conditions - **DC5: The lighting conditions should be stable and optimized to ensure effective makeup removal visualization.**

*3.4.3 Insufficient Awareness regarding Facial Painting's Harmfulness.* Xiqu performers often have demanding schedules and need to take care of their physical well-being. Most of our interviewees (N=12) have reported that they have experienced skin problems or discomfort due to performing since they are too exhausted to remove makeup promptly, leading to unnecessarily long periods of wearing makeup, which in turn triggers various skin issues. (P8: "Sometimes I'm so tired that I don't realize I've been wearing makeup for too long. If I become aware that I've been wearing makeup excessively for several days, I'll definitely make sure to remove it promptly."). By visualizing the trend of time spent wearing makeup, performers can become more aware of how long they have their makeup on.

This awareness aids in timely makeup removal, better skin health management, and prevention of issues like irritation or clogged pores - **DC6: The prototype should provide some indication of wearing duration to enhance user awareness and interaction.**

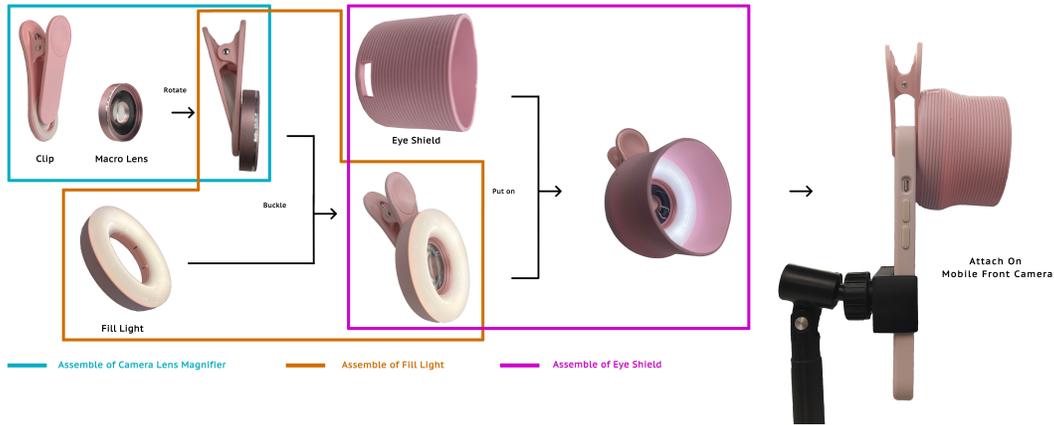
## 4 PROTOTYPE DESIGN

With the guidance of six design considerations, we proposed *EyeVis*, an interactive makeup residue visualization system for Xiqu performers. *EyeVis* contains two parts: (1) a fill light & camera lens magnifier & eye shields for capturing required images and (2) a mobile app for visualization and user control. We provide the design details and alignments with design considerations in this section.

### 4.1 Structural Design

To design a portable, low-cost, and easy-of-use prototype (**DC4**), we consider utilizing the camera of mobile phones with low-cost adjustment to match the requirements of capturing images (shown in figure 2). One of the best devices to capture the eye area is the eye tracker [24], however, it is not applicable when removing makeup because the whole face needs to be cleaned. There are also some existing interactive applications (e.g. *VISIA* [27] and *You Look Good Today* [15]) that capture the whole face for skin imaging analysis, however, systems like *VISIA* have a huge hardware structure that is too cumbersome for the user to move as a portable device. Although systems similar to *You Look Good Today* utilize mobile cameras for skin analysis, which is portable, the performance of the algorithms highly depends on the illumination conditions and the image quality. Therefore, to focus on the eye area instead of the whole face (**DC1**), we consider utilizing a mobile camera lens magnifier (cyan area in Figure 2) to make sure the images that mobile phones take to satisfy a clear range of focal lengths for mainstream mobile models, and we use a fill light (orange area in Figure 2) with eye shields (purple area in Figure 2) so that users can place eye shields against the skin and block out external light sources, relying only on the fill light source to take the picture, which ensures a constant illumination effect (**DC5**).

*4.1.1 Camera Lens Magnifier.* Based on formative study results, we aim to design a system that the camera can capture around the eye area (**DC1**) instead of the whole face. However, the focal length of the camera varies greatly from one mobile phone model to another. In addition, if the full-face shooting mode is used and the eye portion is positioned, on the one hand, it is necessary to deal with the special case of, for example, taking a blurred picture, and on the other hand, the clarity of the cropped picture is not guaranteed, which may lead to a decrease in the accuracy of the image processing algorithm. Additionally, although some mobile devices have a macro lens, it is usually applied in the back, instead of the front. Therefore, we designed a camera lens magnifier, which physically changes the mobile camera lens by applying a clip on the front camera of the mobile phone. We choose to apply a lanthanide optical glass macro lens [2] with a 3x effect to achieve the magnification without losing image qualities. The clip is made of ABS plastic. The clip and macro lens can be assembled together by rotating the circle (cyan area in figure 2). With the camera lens magnifier, users can take higher quality pictures from a closer distance than with a mobile phone lens only.



**Figure 2: Structural Prototype Assemble: Camera Lens Magnifier (in cyan), Fill Light (in orange), and Eye Shields (in purple)**

**4.1.2 Fill Light.** To achieve a stable illumination condition (DC5) when taking photos, we believe that blocking external light sources and generating light spontaneously is the most appropriate approach. For light generation, we use a rounded fill light that can be buckled by the camera lens magnifier (orange area in figure 2). The advantage of the circular design is that we allow the built-in lamp beads to be evenly spaced inside the circle, resulting in a stable and well-positioned light source. Since it takes energy for the light beads to emit light, to enhance portability (DC4), we also set up a battery inside so that the fill light can be recharged and used.

**4.1.3 Eye Shields.** The eye shields are designed for two purposes: (1) blocking the environment light for achieving stable illumination (DC5) and (2) making sure the images are taken at a constant distance (DC1 & DC2). We utilized a skin-friendly material, silicone [8, 34], to build the eye shields as a circular shape and it can be put on the fill light (purple area in figure 2). Based on the camera lens with magnifier, we calculated a standard focus length as the height of the eye shields, and we set the diameter of the eye shields as 7 centimeters, which suits both fill light and human eyes.

## 4.2 Mobile App & Algorithm Design

As we utilized the mobile app's camera as a part of our system, it is natural and convenient for users to interact with a mobile app. We designed 3 workflows for Xiqu performers to help them to visualize their makeup residue (DC2 & DC3) and record their makeup-wearing time (DC6), which are (1) take photos of eyes without makeup, (2) timing & visualization, and (3) time trend. The details are shown in Figure 3. For the usage, users follow these steps: (1) take eye photos without makeup as a baseline, (2) record the makeup wearing time before removing makeup, (3) remove their makeup, (4) take a photo again to visualize makeup residue, and (5) see the usage record and makeup wearing time trend.

We also embedded three vision-based algorithms inside the app, which are (1) eye feature point localization, (2) hsv-uv simulation, and (3) binary threshold, to achieve the visualization effects of our system (shown in Figure 5).

**4.2.1 Prerequisite: Eye feature points localization.** Even though we considered a circular design that fits close to the eyes when designing the structure, there are still cases where the eyes are not centered as each user's face shape is not exactly the same. And the size of the picture taken will be different due to different mobile phone models. To unify this situation, we developed an algorithm of eye feature point localization for each image and resized it to the same size. There are some existing works related to facial feature point localization that achieve high accuracy [44, 79, 85], however, none of them focuses on the eye area only. To enhance the portability of the deployment, we referred to the face feature points localization and iris landmark detection methods (FaceMesh) [72] on Google MediaPipe [47]. Based on their pre-trained model, we can detect 468 facial landmarks with particular regions (we use a sample human face image, which is provided by the user, to run this process). Next, we collected two index arrays with 16 elements, which represent the left & right eye areas, and found the bounding box. Then, we replaced the bounding areas with our image input and run the FaceMesh process again. Finally, we can get the localization feature points and map to the original input (setting bounding boxes and paddings). The pipeline details are shown in Figure 4.

**4.2.2 Workflow1: Take photos of eyes without makeup.** To let users compare w/o (with / without) makeup images (DC2), we designed the workflow of letting users take pictures of no makeup first. In this step, if the user first opens the app, then the user needs to take a photo of the whole face as input for the prerequisite algorithms. Next, the user needs to take two pictures, one with eyes open and one with eyes closed (shown in Figure 3, Workflow1), so that when taking subsequent pictures with makeup if the user doesn't pay attention to the eyes open or closed, the user will still have a similar picture without make-up to match for better comparison. The matching scheme is to calculate the distance of the highest / lowest eye feature points to determine the degree of opening and closing of the eyes.

**4.2.3 Workflow2: Timing & Visualization.**

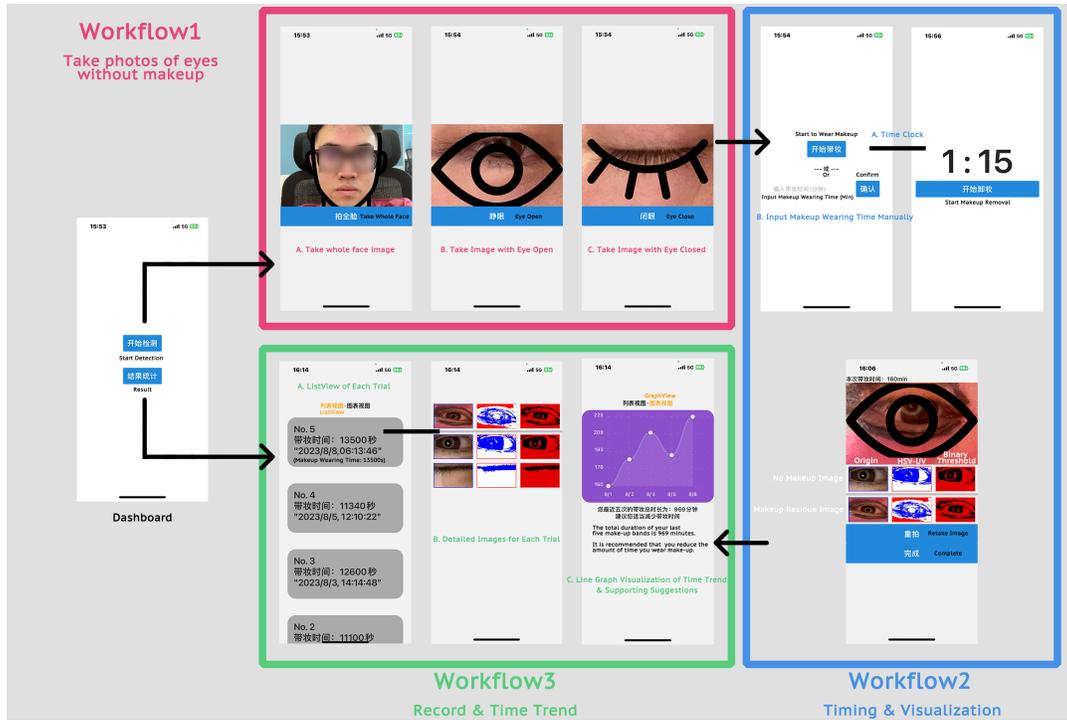


Figure 3: Mobile App Workflow: (1) Take photos of eyes without makeup (whole face, open and close eyes), (2) Timing and visualization (HSV-UV simulation and Binary threshold, comparison w / o makeup), (3) Time trend and record display.

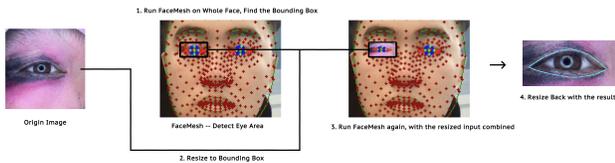


Figure 4: Sample pipeline of Eye feature points localization. (1) Run Facemesh on whole face as a reference to find the bounding box of eye feature points, (2) Resize the origin image to the bounding box size and replace it with the eye area of the whole face, (3) Run Facemesh again with the input combined image, (4) Extract the eye feature points and resize back to the final result.

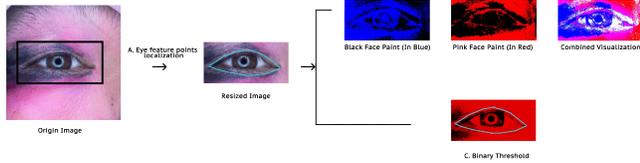
**Timing.** After taking photos without makeup, the user starts to wear the makeup with the face paint. To record the makeup-wearing time (DC6), we designed a timing scheme. For ease of use, the user can choose to record the wearing time by a time clock or input the wearing time manually (shown in Figure 3, Workflow2, A & B).

**Visualization: HSV-UV Simulation.** To augment the black and pink areas of residual face paint (DC3), we first tried special light sources [50]. Ultra-Violet (UV) light has been widely used in skin disease diagnosis, and it is also effective in allowing cosmetic residues

to fluoresce [9]. However, the eye area can be irreversibly and severely damaged by exposure to UV light [48]. Therefore, we try to make the UV simulation by utilizing color spaces and image processing & filtering. Although our system is designed for stable illumination, the RGB color space is sensitive to very small changes in brightness (R, G, B values can vary greatly). Therefore, we utilized HSV color space to filter similar colors of black and pink face paints. HSV color filtering contains these steps: (1) enhance the intensity of the blue channel, (2) change from RGB to HSV color space, then set HSV thresholds (lower and upper bound) of black and pink face paints, (3) apply HSV filtering and transfer to single-color images (blue for black face paints, and red for pink face paints). Figure 5, B, shows a sample of the algorithm effect, the blue image means black face paints are detected, while the red image links with pink face paint. The pink color in the combined image shows the intersection of black and pink face paint.

**Visualization: Binary Threshold.** To better describe the situation around the eyes, we designed another image-processing algorithm to visualize the edges & contours of the eye-around area. The idea is to transfer the color space of the image from RGB to gray, then we set a lower bound and upper bound for thresholding: pixels inside the bound will be marked as black, otherwise white. Finally, we apply edge detection and give the output image a striking color map to replace the white area. Figure 5, C, shows a sample of the algorithm effect, the user can see edges & contours of eyeliner around the eye area.

Finally, there are two rows of image sets, which represent the w/o makeup images. For each row, the original image, HSV-UV simulation image, and binary threshold image are shown as a set for the user to check and compare.



**Figure 5: Overview of Vision-based Algorithms: firstly run (A) eye feature points localization (clip and resize the image based on the bounding box), then run (B) HSV-UV simulation (we provide blue color for black face paint, and red color for pink face paint visualization) and (C) Binary Threshold to visualize the makeup residual.**

**4.2.4 Workflow3: Time Trend.** After recording makeup wearing time for each time, the user can see a time trend visualization (DC6) provided by *EyeVis*. We provided 2 types of record visualization: (1) a list view of each use, the user can click into each use for rechecking the images captured for each use, and (2) a line graph to represent the time trend of the last five makeup experiences (shown in Figure 3, Workflow 3, A & C).

### 4.3 Implementation

**4.3.1 Algorithm Development.** To localize the eye feature points, we used the MediaPipe [47] Python library (version 0.9.1, which supports face landmark detection) and Numpy (version 1.24.3) for image cropping and resizing. We also utilized the OpenCV [11] Python library (version 4.8.0.74) for HSV-UV simulation and Binary Threshold visualization. Finally, we integrated all code scripts into the back end of the mobile app.

**4.3.2 Mobile App Development.** To make our app accessible to more users, we use a cross-end framework, react-native [21]<sup>2</sup>, to develop our mobile app in JavaScript scope. For algorithm deployment, we developed a back-end by using FastAPI of Python<sup>3</sup> and deployed the script to a Tencent Cloud<sup>4</sup> remote server (running on Linux, CentOS 7.9). Then, we deploy our code to both XCode and Android Studio for iOS and Android application packages. Finally, we release our system by using an online app distribution service called Pgyer<sup>5</sup>.

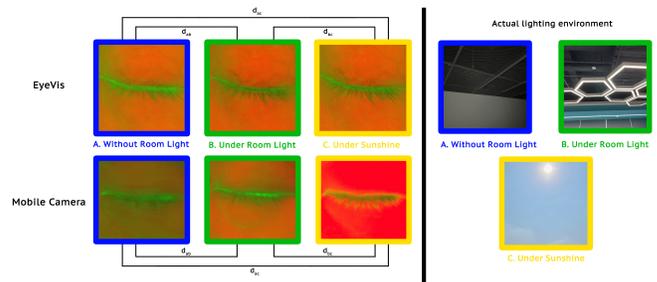
## 5 TECHNICAL EVALUATION

To evaluate the robustness of the algorithms, we used *EyeVis* to collect a testing database that contains 259 eye images from four co-authors. The database collection procedures include: (1) take eye pictures with no makeup, (2) take eye pictures with full makeup (with pink and black face paint), and (3) take eye pictures with

makeup residue (i.e. remove the face paints in parts of the eye-around area but not all). To extend the generalizability of the dataset, we collected images that contain both open and closed eyes, and the number of images in different categories is evenly distributed. Although there are many large-scale models that are gradually being applied to image segmentation [37, 46], their effectiveness is limited to recognizing more common objects, and they are generally ineffective for images with only eyes and makeup. Therefore, all images are manually labeled with makeup areas (black & pink face paints and eye-around contours). Technically, we use Computer Vision Annotation Tool (CVAT [17]) to manually label 3 classes (pink face paint, black face paint, eye) with polygons for each image.

### 5.1 Illumination Stability

To validate the stability of the illumination of *EyeVis*, we tested 10 groups of data, each group contains a series of images that are taken by the same eye with the same camera settings ( $ISO = 200$ ,  $Shutter = \frac{1}{50}s$ ,  $Focal\ Length(f) = 26mm$ ,  $Aperture = \frac{f}{2.2}$ ,  $White\ Balance = 5200K$ ), in various lighting surrounding conditions: (1) without room light, (2) under room light, and (3) under sunshine. Then we transformed the images in HSV color space, and we calculated the Euclidean distance [19] ( $d = \sqrt{d_h^2 + d_s^2 + d_v^2}$ ) of each two images ( $img_0, img_1$ ) in each pixel in terms of hue ( $d_h = \min(abs(h_1 - h_0), 360 - abs(h_1 - h_0))/180$ ), saturation ( $d_s = abs(s_1 - s_0)$ ), and value ( $d_v = abs(v_1 - v_0)/255$ ). Figure 6 shows a sample of the grouped data, compared with images taken by a mobile phone’s front camera directly at the same distance as using *EyeVis* (we used iPhone 13’s camera for testing). Table 1 shows our test average results, based on the average data, we observed that all Euclidean distances of the two chosen images are significantly lower than taking images by mobile directly. Therefore, we can conclude our system achieves a stable illumination environment.



**Figure 6: LEFT: Sample Illumination Test: w / o EyeVis in HSV color space. RIGHT: Actual lighting surrounding environments. The images are taken under 3 different lighting conditions: a) without room light, b) under room light, and c) under sunshine.**

### 5.2 Visualization Methods

We evaluate the robustness of visualization methods: 1) eye feature points localization, 2) hsv-uv simulation, and 3) binary threshold, by comparing the regional overlap between algorithm results and

<sup>2</sup><https://reactnative.dev/>

<sup>3</sup><https://fastapi.tiangolo.com/>

<sup>4</sup><https://www.tencentcloud.com/>

<sup>5</sup><https://www.pgyer.com/>

	Avg $d_{ab}$	Avg $d_{ac}$	Avg $d_{bc}$
EyeVis	0.86	0.73	0.69
Mobile Camera	1.91	4.60	3.09

**Table 1: Average Euclidean Distance Comparison Under Different Lighting Conditions: w / o EyeVis**

manually labeled area. We utilized the Shapely [26] Python library to do the polygon-area transformation and overlap calculation.

**5.2.1 Eye Feature Points Localization.** To evaluate the accuracy of eye feature point localization, we used our manually marked eye area to compare with the algorithm results, and we define  $a_1 =$  *algorithm results inside manually marked eye area*,  $a_2 =$  *manually marked eye area*,  $r_{eye} = \frac{a_1}{a_2}$  as the quantitative value that indicates the accuracy (see Figure 7, C, for the region comparison). The areas of the algorithm results are formed by polylines of the 16 feature points.

**5.2.2 HSV-UV Simulation.** For the evaluation of hsv-uv simulation, we also define  $a_3 =$  *algorithm results inside manually marked pink area*,  $a_4 =$  *manually marked pink area*,  $r_{pink} = \frac{a_3}{a_4}$  and  $a_5 =$  *algorithm results inside manually marked black area*,  $a_6 =$  *manually marked black area*,  $r_{black} = \frac{a_5}{a_6}$  as quantitative indicators to justify the accuracy (see Figure 7, D & E).

**5.2.3 Binary Threshold.** To test the robustness of binary threshold effects, we examined the area outside of the manually delineated eye range, and if the pixel labeled black in the threshold range happened to fall within the manually delineated black paint area, the pixel would be labeled as "success". We define  $n_1 =$  *number of inside-threshold pixels marked as "success"* (outside-threshold pixels marked as "not success"),  $n_2 =$  *number of inside-threshold pixels outside eye area*,  $r_{bin} = \frac{n_1}{n_2}$  as the quantitative indicator to justify the accuracy of the binary threshold (see Figure 7, F).

	$r_{eye}$	$r_{pink}$	$r_{black}$	$r_{bin}$
Avg Overlap Rate	96.1%	80.4%	85.2%	82.8%

**Table 2: Average Overlap Rate among eye feature points localization ( $r_{eye}$ ), HSV-UV simulation ( $r_{pink}$  and  $r_{black}$ ), and Binary Threshold ( $r_{bin}$ ).**

Table 2 shows the testing results among 259 images in our self-maintained dataset.  $r_{eyes}$  achieves the highest average overlap rate of 96.1%, which is a testament to the robustness of the algorithm on the task of locating the eyes. For the HSV-UV simulation method, the results for the pink area ( $r_{pink}$ ) are slightly lower than the black area ( $r_{black}$ ), and we speculate that this may be due to the fact that each person applies oil colors in different shades, and thus the HSV may not have covered all the cases when setting up the color filter. For  $r_{bin}$ , some images are affected by things other than the eye, such as eyebrows and other objects that resemble a black border, potentially causing some pixels to be "not successful". However, the average overlap rate for each indicator still remained above 80%. As we developed our algorithms without machine learning to reduce the burden of back-end computing, such an overlap rate

indicates our algorithms can achieve considerable effects, and are ready to be deployed and tested by users.

## 6 DEPLOYMENT STUDY

To examine the usability and effectiveness of *EyeVis*, we ran a 7-day deployment study with 12 participants. Based on the study results, we discussed and explored potential design implications for future work. To our knowledge, it is the first time a prototype of this kind has been created and tested in a long-term deployment study.

### 6.1 Participants & Apparatus

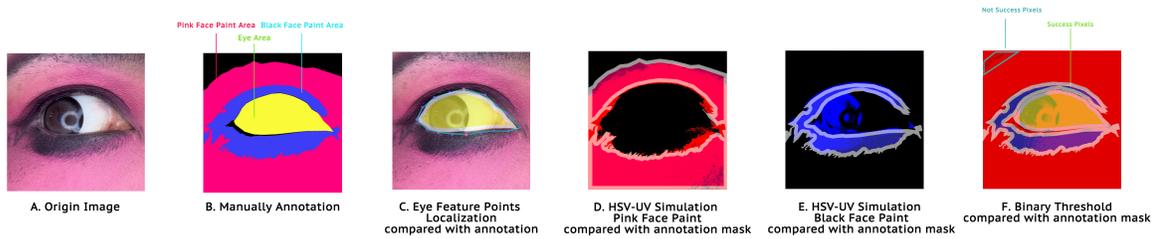
We recruited 12 participants (3 males, and 9 females, aged 16-55) by posting recruitment messages on the online community. All participants are from Xiqu troupes or academies and have experience in Xiqu makeup. Before the experiment, each participant signed a consent form (approved by the university's ethical review) to agree to collect user data. As we deployed our back-end in a remote server, and participants can use the prototype anywhere on any Android or iOS device. For remote users, we also sent a package of fill lights, camera lens magnifier, and eye shields to each of them. 100 RMB was given to each participant as the reimbursement.

### 6.2 Task Design & Procedure

We designed a 7-day user study to explore the usability and effectiveness of our system. Before the study officially starts, we record eye images (by using the same camera settings as *EyeVis* without letting participants see the visualization and remove accordingly) in a daily makeup removal process for each participant, so that we can perform the baseline (daily makeup removal) & *EyeVis* comparison. For Day 1, we have an introduction session for kicking off, we introduce users to (1) how the structure part is installed, (2) how the software part is installed, and 3) how to use the app for the makeup removal experience. Then, participants went through three workflows of the mobile app. Before taking photos without makeup, participants were asked to clean their faces carefully, to avoid residue before applying make-up. Then participants wear eye makeup and remove it for a test. Participants were also requested to complete a NASA TLX form [29] to document their daily makeup removal routine experience, serving as a baseline data recording. After the introduction session, for Day 2 to Day 7, participants were asked to repeat the makeup and makeup removal process at least 4 more times (i.e. a total of at least 5 times of using) during performance or daily practice, to examine the effects of makeup-wearing time trend. To simulate the most realistic use, participants were given the opportunity to use their personal favorite make-up removers to remove their make-up. For each use, we recorded the photos that participants took and the makeup removal time in the back end for analysis, and the consent of the participants was obtained. Finally, on Day 7, participants were asked to fill out a NASA TLX form and a 7-point Likert scale questionnaire to evaluate the usability, comfortability, and effectiveness of the system [31], and a final semi-structured interview was conducted.

### 6.3 Data Analysis

Four co-authors open-coded the deployment study data in terms of qualitative statistics (user feedback in semi-structured interviews,



**Figure 7: Sample Technical Evaluation of Visualization Methods: A shows a sample origin image; B shows the related manually annotated image; C, D, and E show the overlap between algorithm results and manually annotated area, in terms of the eye area, pink face paint area, and black face paint area, respectively; F shows “success” points (inside black area) in the binary threshold image.**

and ratings in the NASA TLX forms, in terms of the usability of the system, and ratings in 7-point Likert scale questionnaires regarding the features of the system) and quantitative statistics (the images contained by using *EyeVis* and daily makeup removal). Then we summarized the data in terms of the system’s usability, the considerable long-term effects of use, and the quantitative comparison between using *EyeVis* and *Daily Makeup Removal*.

For quantitative comparison of makeup residue, we take the following approaches:

- (1) Run the *EyeVis* back-end algorithm once with the captured daily makeup removal image as input, we can get the area where the pink face paint and black face paint are located, and calculate the ratio of the area it occupies to the whole image area ( $r_{p\_baseline} = \frac{\text{pink face paint area}}{\text{whole baseline image area}}, r_{b\_baseline} = \frac{\text{black face paint area}}{\text{whole baseline image area}}$ ).
- (2) We can get the area of pink and black face paint in the eye image captured by the user every time the participant uses *EyeVis*, and calculate the ratio of the area to the whole image ( $r_{p\_EyeVis} = \frac{\text{pink face paint area}}{\text{whole EyeVis image area}}, r_{b\_EyeVis} = \frac{\text{black face paint area}}{\text{whole EyeVis image area}}$ ).
- (3) Get the mean value of the data ( $r_{p\_EyeVis\_mean} = \frac{r_{p\_EyeVis}}{5}$ ,  $r_{b\_EyeVis\_mean} = \frac{r_{b\_EyeVis}}{5}$ ) when the user uses the data five times and compare this mean value with *EyeVis*.

Note that we don’t use the makeup removal time as an indicator of how effective the system is compared to daily makeup removal, because when using the *EyeVis* system, users tend to double-check to see if they’re getting it off cleanly, which can lead to longer make-up removal times. Similarly, the time taken to remove makeup depends on the thickness and complexity of the makeup. Therefore, we only use “how completely the makeup is removed” to judge the system’s effectiveness.

## 7 RESULT

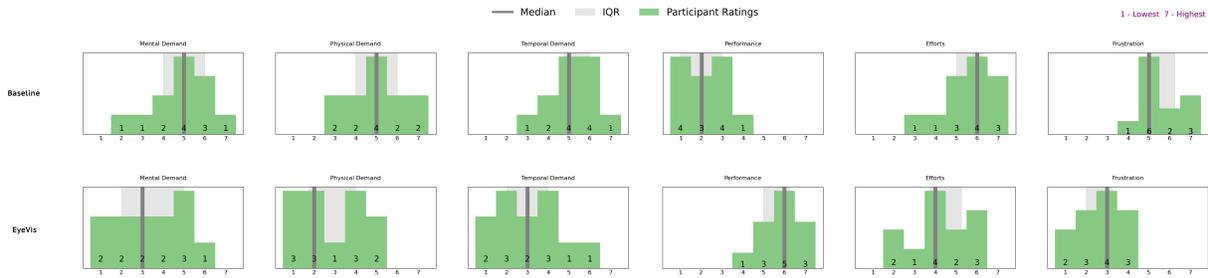
### 7.1 System Usability

**7.1.1 Overall Usage of the System.** Overall, all participants indicated that the usage of an assistant makeup removal system was helpful for them to remove makeup more effectively. The system’s

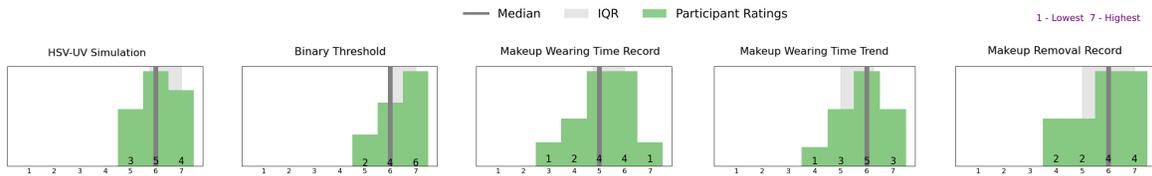
usability was evaluated by the ratings of the NASA TLX questionnaire. In terms of performance (makeup removal completeness), they agreed that compared with baseline (makeup removal without *EyeVis*, recorded before introducing *EyeVis* to participants) ( $Md=2, IQR=2$ ), with the assistance of *EyeVis*, they can remove makeup more completely ( $Md=6, IQR=1.25$ ). Participants also appreciated the *EyeVis* system, noting its significant assistance in reducing the efforts ( $Md=4, IQR=1.5$ ) and frustration ( $Md=3, IQR=1.25$ ) with the makeup removal process. They found that utilizing the assistive system reduced the physical ( $Md=2, IQR=2.25$ ), mental ( $Md=3, IQR=3$ ) and temporal ( $Md=3, IQR=2$ ) demands involved in removing their makeup, especially when compared to their daily routine, which often required repeated wiping to ensure that all the makeup was properly removed. “With the *EyeVis* system, I can easily determine when to stop wiping, avoiding over-cleansing and discomfort due to excessive rubbing.”-P3.

**7.1.2 Visualization methods.** The ratings of the 7-point Likert scale questionnaire evaluated different visualization methods of the system (shown in Figure 9). *EyeVis* provides users with two visualization methods, HSV-UV Simulation and Binary Threshold, to identify makeup residue. There are no significant advantages or disadvantages between **HSV-UV Simulation** ( $Md=6, IQR=1.25$ ) and **Binary Threshold** ( $Md=6.5, IQR=1$ ) based on the ratings. Users, on average, expressed their appreciation for both techniques based on their preferences. However, most participants ( $N=10$ ) did appreciate specific aspects of Binary Threshold, as P10 stated, “I prefer Binary Threshold because it creates a more intuitive color contrast around eye area”. In contrast, limited participants preferred HSV-UV Simulation (P1, P6), saying, “I mostly focus on the HSV-UV visualization as I want to quickly identify and remove the black & pink face paint inside the eye”. Additionally, users utilized the visualization approach to monitor different aspects of makeup residue, ranging from overall residual makeup to more detailed areas, “HSV-UV helps me gauge the extent of general residue, while Binary allows me to examine the residue in more detail”-P7.

**7.1.3 Objective Feedback of Time Trend.** Participants highly valued the *EyeVis* system’s makeup-wearing time record ( $Md=5, IQR=1.25$ ) and makeup-wearing time trend ( $Md=6, IQR=1.25$ ) as the most crucial aspect of the system. They found the time recording intuitive and user-friendly, P4 expressed, “I simply click on the interface at



**Figure 8: Participants' ratings (regarding the usability of the system) on a scale of 1-7 using NASA Task Load Index. Top: ratings for daily makeup removal. Bottom: ratings for makeup removal by using EyeVis.**



**Figure 9: Participants' ratings (regarding the feature of the system) on a 7-point Likert scale questionnaire. For the performance scale, a higher rating means better performance.**

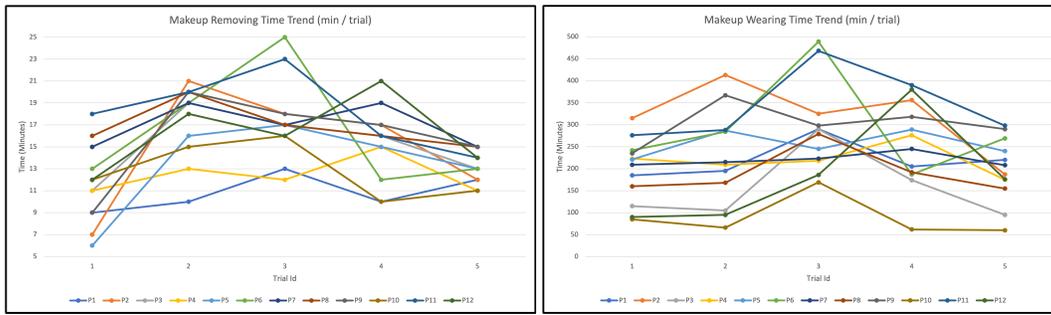
the start and end of my makeup process, and it automatically records the time. This allows me to keep track of my skin's condition more accurately". The makeup-wearing time trend further enhanced participants' awareness. P7 shared, "After wearing makeup for three days, I checked the trend feature and realized that I had been applying makeup too frequently, which could lead to skin problems. Previously, I might have ignored or not remembered this. So, I followed a meticulous skincare routine". Some participants also mentioned additional benefits of viewing recordings, which helped them monitor their skin condition more precisely. "Occasionally, I compared my eye pictures from a week ago with today's and noticed that my dark circles had worsened. It made me realize I should take some time off" -P10. We also observed that the time trend is an initial approach in skincare recommendations, as P7 stated: "It's nice to be able to visualize how long I can wear makeup, but it would be nicer to recommend skincare products or skincare regimens based on my actual needs."

## 7.2 Persistent Effects

**7.2.1 Consistent Trend of Makeup Removal Time.** Figure 10 shows a general trend in makeup removal time for 12 participants. As each participant has different scenes with make-up (rehearsal or performance) and different make-up removal habits, we can see that the initial values in the graphs are quite different, but we can still find some patterns in them: First, during the initial usage of the system, participants received assistance from the researchers. However, as they continued interacting with the system over time, their interest in other features gradually increased. As a result, the second usage of the system showed a slight upward trend. Participants took a casual approach to makeup removal initially, which resulted in a relatively short time spent removing makeup. Furthermore,

as they progressed and understood the importance of removing makeup thoroughly the duration of the procedure trended upwards. As the participants continued to use the assistant system during the makeup removal process, they gradually became more familiar with its functionality and developed a routine. This familiarity and routine led to increased efficiency in removing makeup, resulting in a decrease in the overall duration of the process. As P5 stated, "At the beginning, I wasn't very familiar with the Usage of EyeVis, but after using it three or four times, I gradually mastered the tricks and it allowed me to remove my makeup more quickly". Figure 10 also shows a positive correlation between the makeup-removing time and makeup-wearing time, especially for P6 and P11. This suggests that participants who wore makeup for longer took more time to remove it completely, "During a long-time performance, makeup needs to be touched up repeatedly, resulting in a thicker layer. Consequently, it takes me more time to remove the makeup."-P12.

**7.2.2 Change in Makeup Removal Practices.** We observed that some participants' makeup removal habits changed after the 7-day experiment, e.g. some participants who had previously paid less attention to the make-up removal process paid more attention to this process: As P9 stated: "Previously I was very casual about removing my makeup and after using EyeVis I would subconsciously check to see if it had been removed completely". For experienced and skin-conscious actors, their make-up removal habits are not subject to much change. P2 mentioned that: "Having been doing Xiqu make-up for many years now, it's a fact that eye make-up is hard to remove. However, since I have more experience, my make-up removal habits have not been affected much." Additionally, there were some participants who, as heavy users of face paint, had increased the importance they placed on removing their makeup: "As a frontline actor, I wear my



**Figure 10: The makeup-wearing & removing time trend among all participants in five trials. LEFT: Makeup removing time trend. Overall, participants enhanced their makeup removing time to familiarize themselves with specific visualization methods of the system, then it reduced when they got used to it. RIGHT: Makeup wearing time trend. Generally, participants enhanced their awareness of makeup removal.**

*make-up for a very long time every day, and by the time I take it off I'm often tired from several performances, so I remove my make-up in a haphazard manner. This system has really helped me to improve my make-up removal habits and hopefully, my skin will slowly improve.*  
- P6

### 7.3 Quantitative Comparison of Makeup Residue: *EyeVis* vs Daily Makeup Removal

We report each participant's ratios of specified regions (pink & black face paints) in Appendix B. Based on our collected data:  $r_{p\_baseline}$  (avg = 19.84%, std = 6.51%),  $r_{p\_EyeVis\_mean}$  (avg = 4%, std = 1.66%),  $r_{b\_baseline}$  (avg = 19.48%, std = 6.81%), and  $r_{b\_EyeVis\_mean}$  (avg = 3.63%, std = 1.61%), we observe that: (1) **Significant Reduction in Face Paint Residual Compared with Baseline.** There's a marked decrease in the percentage of both pink and black paint areas when measured using the *EyeVis* system compared to the baseline. This could suggest that the *EyeVis* system is more efficient or conservative in removing the makeup. (2) **Consistency and Variability.** The *EyeVis* system measured two colors with a small standard deviation, indicating better consistency. This means that the *EyeVis* system standardized the makeup removal process more effectively than the baseline method (daily makeup removal), resulting in more consistent results across participants. (3) **Diversity in Individual Cases.** Even though the system shows greater consistency, there is still a range of responses among participants. This diversity might be attributed to individual differences, such as variations in face size or shape, which could affect the makeup removal completeness.

## 8 DISCUSSION

In this section, we summarize the most important findings, observations, and limitations of our research. We also highlight a few design takeaways from the research and potential future directions.

### 8.1 Key Contributions

We investigated the practices and challenges encountered by Chinese traditional opera (Xiqu) performers in relation to skin condition issues, by conducting an online survey and semi-structured

interviews. Specifically, we identified that incomplete makeup removal is the key cause due to human factors, and designing an interactive makeup residue visualization application might be helpful for Xiqu performers in aid of their skincare. From there we derived six design considerations in section 3.4.

Then, we implemented these design considerations in a prototype called *EyeVis* with these parts: mobile app + fill light + camera lens magnifier + eye shields, and in the visualization part, we embedded image processing algorithms including (1) eye feature points localization, (2) hsv-uv simulation, and (3) binary threshold. We conducted a technical evaluation for illumination stability and visualization methods by utilizing a self-maintained eye image database. We also evaluated the usability and long-term effects of *EyeVis* by running a 7-day deployment study with 12 participants. We arrived at the following key conclusions and insights to guide future design: (1) Two Visualisation methods enhance the color display of makeup residue to help users confirm that it has been removed, (2) *EyeVis* is designed to meet the portability requirements of users, (3) Long-term use of *EyeVis* helps to improve the awareness of make-up removal and skincare.

### 8.2 Algorithm Robustness Improvement & Customization

To reduce the burden on the back-end server, we do not use the mode of machine learning model training for image processing, but also as a result the visual algorithms will have the problem of weak generalization performance. There are various makeup datasets with whole faces (e.g. [42, 86]) that have been used in vision algorithm development, however, to our knowledge, there are no datasets using *EyeVis*, a camera with a special focal length that produces images, have been collected and widely used. Therefore, increasing the diversity of the dataset can also provide more possibilities to improve the robustness of the algorithm. For example, (1) collect more images of Xiqu performers wearing eye makeup with multimodal attributes (e.g. correspondence between eye make-up and Xiqu acting roles, the degree of opening and closing of the eyes, etc.), and (2) invite experts, such as dermatologists, to perform large-scale and more precise image annotation (e.g. distinguish

black face paint and eyelash, distinguish different levels of residue, etc.).

Regarding algorithm customization, we note that it is necessary to study specific operatic characters and develop individualized visualization methods. Even though pink and black face paint is the more common makeup colors, Xiqu performers often adjust the intensity of their makeup according to the stage lighting or the special requirements of their roles, and some local operas use special face paints (e.g. Yingge [28], Tibetan Opera [12]) that require makeup with multiple complex colors. Generalizing the visual algorithms we have designed to more niche roles and genres will benefit more Xiqu performers.

### 8.3 System Portability

With the development of mobile phone chips, the response speed of its CPU can also meet the requirements of using mobile phone cameras as hardware for eye-capture tasks has become more common (e.g. (1) eye-tracking in virtual reality [60], (2) eyelid detection [23], etc.). However, many systems related to skin imaging have yet to apply ubiquitous computing to mobile phones. Compared to professional skin imaging devices for medical use (e.g. VISIA [27], which contains a desktop operating system with a huge casing that covers the whole face), mobile phones are available to everyone, and the fill light and lens design of *EyeVis* can be taken out and used anywhere and does not take up much space.

Despite this, *EyeVis* is still a stop-gap measure. As we utilized mobile cameras as image-capturing tools, even if we use a camera lens magnifier, the photos taken by different models of mobile phone cameras will still be slightly different. The choice between building our own camera module to achieve image uniformity across users, or utilizing the computing resources of the phone to mitigate the development costs is a trade-off. In addition, since each user's facial features are different, it would be possible to personalize our system even more if the eye shields could all be cut to fit different people's faces. How to reduce the development cost as much as possible while achieving the unity of the system use effect, will be worth studying in future work.

### 8.4 Generalizability

**8.4.1 Wider Area: From Eye to Skin.** Since the launch of the formative study, we focused on the eye area, but in reality, our system can be applied to any skin area other than the eye (e.g. nose, lips). As long as the area can be completely covered by the eye shield, we can still provide stable illumination and high-dimension image quality for image processing. Even if complete coverage by the eye shield is hard, we can fine-tune the focus of the macro camera magnifier and the diameter size of the eye shield to achieve the desired effect.

**8.4.2 Wider Audiences & Scenarios: From Xiqu Performers to People Who Makeup Frequently.** Our system is not limited to detecting face paints that are used by Xiqu performers. For everyday makeup & skincare products, some people who makeup frequently also face problems with incomplete removal or wearing. For example, sunscreen, people may not be sure if it has been fully applied to the designated area. While sunscreen has effects to reflect ultraviolet light [52, 54], the idea of uv simulation can also be applied to detect the level of sunscreen coverage. Therefore, our system can also be

applied to everyday makeup & skincare products and serve people who makeup frequently.

**8.4.3 Wider Usage: Skin Data Collection Tool.** In addition to visualizing and comparing makeup residue, our system can also be used as a skin data collection tool. Skin image databases are used for various scenarios, including clinical diagnosis [40, 64], skin condition prediction [70], racial disparities analysis [7], etc. Compared to existing skin data collection tools, our system has these advantages: (1) Portable and easy to use, we do not need another customized camera or hardware, and one mobile camera is enough for achieving effects under macro cameras, (2) Cost-effective in centralizing data, unlike devices that collect images locally, our system is directly connected to the remote server, so the remote central server has direct access to the images collected by the systems in each location. Therefore, our system can provide skin imaging developers with the convenience of collecting data and provide some reference values for the advancement of related technologies.

## 9 CONCLUSION AND FUTURE WORK

This work emphasized the skin health issues Xiqu performers experience as a result of applying heavy metal-containing face paints for an extended period of time. To investigate the practices and challenges that Xiqu performers face during the skincare process, we conducted an online survey and semi-structured interviews. The findings showed that human-induced skin issues for Xiqu performers were primarily caused by improper makeup removal, particularly from the eyes. To this end, we proposed *EyeVis*, a prototype that can visualize the makeup residue around the eye area and keep tracking the makeup-wearing time of Xiqu performers. The seven-day deployment study verified the usability and long-term results of *EyeVis*. The findings showed that *EyeVis* increases Xiqu performers' awareness of skincare and makeup removal while also boosting their self-assurance and security in skincare.

In addition to highlighting the significance of proper makeup removal and skincare in the performing arts industry, we also offered insights into the skin health issues experienced by Xiqu performers. For future work, a larger deployment study to validate the effectiveness of *EyeVis* in a larger sample of Xiqu performers will be conducted, and additional features such as personalized skincare recommendations based on skin type and makeup usage will be incorporated into the system. In conclusion, the work offers a promising direction for future investigation into skincare and makeup removal in the performing arts sector and helps to promote and preserve the intangible cultural heritage of practitioners.

## ACKNOWLEDGMENTS

We thank our reviewers for their constructive feedback and Xiqu performers from Guangdong Cantonese Opera Theatre, Guangzhou Youth Peking Opera Association, Guangdong Dance and Drama College, National Academy of Chinese Theatre Arts, Northern Kunqu Opera Theatre, Suzhou Kunqu Opera Theatre, Changzhou Wuxi Opera Theatre, Sichuan Opera Theatre, and Wenzhou Yue Opera Theatre, for their participation. This work is partially supported by 1) 2024 Guangzhou Science and Technology Program City-University Joint Funding Project (PI: Mingming Fan); 2) 2023 Guangzhou Science and Technology Program City-University Joint

Funding Project (Project No. 2023A03J0001) ; 3) Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (No.2023B1212010007).

## REFERENCES

- [1] Taleb Alashkar, Songyao Jiang, Shuyang Wang, and Yun Fu. 2017. Examples-rules guided deep neural network for makeup recommendation. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 31.
- [2] RO Alekseev, NA Romanov, VI Savinkov, NN Klimenko, and VN Sigaev. 2021. Multicomponent optical glasses with high refractive index. *Glass and Ceramics* 78 (2021), 3–7.
- [3] Onur Asan and Avishek Choudhury. 2021. Research trends in artificial intelligence applications in human factors health care: mapping review. *JMIR human factors* 8, 2 (2021), e28236.
- [4] China Theater Association. 2019. Current development of professional folk theatre troupes. [https://www.chinatheatre.org.cn/jxhd/qgfc/201905/t20190507\\_444010.html](https://www.chinatheatre.org.cn/jxhd/qgfc/201905/t20190507_444010.html). (Accessed on 07/15/2023).
- [5] Tamara Attard and Everaldo Attard. 2022. Heavy metals in cosmetics. (2022).
- [6] İlke Bahçeci Şimşek and Can Şırolu. 2021. Analysis of surgical outcome after upper eyelid surgery by computer vision algorithm using face and facial landmark detection. *Graefe's Archive for Clinical and Experimental Ophthalmology* 259, 10 (2021), 3119–3125.
- [7] Teanna Barrett, Quanze Chen, and Amy Zhang. 2023. Skin Deep: Investigating Subjectivity in Skin Tone Annotations for Computer Vision Benchmark Datasets. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (Chicago, IL, USA) (FAccT '23). Association for Computing Machinery, New York, NY, USA, 1757–1771. <https://doi.org/10.1145/3593013.3594114>
- [8] Jeanette Berry, Pat Black, Rory Smith, and Barbara Stuchfield. 2007. Assessing the value of silicone and hydrocolloid products in stoma care. *British journal of nursing* 16, 13 (2007), 778–788.
- [9] Karina Bispo-dos Santos, Priscilla P Barbosa, Fabiana Granja, Matheus Cavalheiro Martini, Camila Flavia Schettino Oliveira, Desiree Cigarán Schuck, Carla Abdo Brohem, Clarice Weis Arns, Sylvio Jorge Hares Junior, Caetano Padiál Sabino, et al. 2021. Ultraviolet germicidal irradiation is effective against SARS-CoV-2 in contaminated makeup powder and lipstick. *Journal of Photochemistry and Photobiology* 8 (2021), 100072.
- [10] Sylwia Borowska and Malgorzata M Brzóska. 2015. Metals in cosmetics: implications for human health. *Journal of applied toxicology* 35, 6 (2015), 551–572.
- [11] Gary Bradski. 2000. The openCV library. *Dr. Dobbs' Journal: Software Tools for the Professional Programmer* 25, 11 (2000), 120–123.
- [12] Marcia S Calkowski. 1991. A day at the Tibetan opera: actualized performance and spectacular discourse. *American ethnologist* 18, 4 (1991), 643–657.
- [13] Elena Chirikhina, Andrey Chirikhin, Sabina Dewsbury-Ennis, Francesco Bianconi, and Perry Xiao. 2021. Skin characterizations by using contact capacitive imaging and high-resolution ultrasound imaging with machine learning algorithms. *Applied Sciences* 11, 18 (2021), 8714.
- [14] Toby Chong, Nolwenn Maudet, Katsuki Harima, and Takeo Igarashi. 2021. Exploring a Makeup Support System for Transgender Passing Based on Automatic Gender Recognition. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 568, 13 pages. <https://doi.org/10.1145/3411764.3445364>
- [15] Hangzhou Yixi Internet Technology Co. 2016. You Look Good Today - Take a photo to measure your skin quality and remove blackheads with scientific skincare. <https://c2h4.org/>. (Accessed on 08/04/2023).
- [16] Catia Contado and Antonella Pagnoni. 2012. A new strategy for pressed powder eye shadow analysis: allergenic metal ion content and particle size distribution. *Science of the total environment* 432 (2012), 173–179.
- [17] CVAT.ai Corporation. 2023. *Computer Vision Annotation Tool (CVAT)*. <https://doi.org/10.5281/zenodo.8239758>
- [18] Richard Corson, James Glavan, and Beverly Gore Norcross. 2019. *Stage makeup*. Routledge.
- [19] Per-Erik Danielsson. 1980. Euclidean distance mapping. *Computer Graphics and image processing* 14, 3 (1980), 227–248.
- [20] Xianghua Ding, Yanqi Jiang, Xiankang Qin, Yunan Chen, Wenqiang Zhang, and Lizhe Qi. 2019. Reading Face, Reading Health: Exploring Face Reading Technologies for Everyday Health. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300435>
- [21] Bonnie Eisenman. 2015. *Learning react native: Building native mobile apps with JavaScript*. " O'Reilly Media, Inc."
- [22] Lisa Eldridge. 2015. *Face paint: The story of makeup*. Abrams.
- [23] Mingming Fan, Zhen Li, and Franklin Mingzhe Li. 2021. Eyelid gestures for people with motor impairments. *Commun. ACM* 65, 1 (2021), 108–115.
- [24] Gregory Funke, Eric Greenlee, Martha Carter, Allen Dukes, Rebecca Brown, and Lauren Menke. 2016. Which eye tracker is right for your research? performance evaluation of several cost variant eye trackers. In *Proceedings of the Human Factors and Ergonomics Society annual meeting*, Vol. 60. SAGE Publications Sage CA: Los Angeles, CA, 1240–1244.
- [25] Gunilla Färm, Ann-Therese Karlberg, and Cajrola Lidén. 1995. Are opera-house artists afflicted with contact allergy to colophony and cosmetics? *Contact Dermatitis* 32, 5 (1995), 273–280. <https://doi.org/10.1111/j.1600-0536.1995.tb00780.x> arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1600-0536.1995.tb00780.x>
- [26] Sean Gillies. 2013. The shapely user manual. URL <https://pypi.org/project/Shapely> (2013).
- [27] Anne Goldsberry, C William Hanke, and Katherine E Hanke. 2014. VISIA system: a possible tool in the cosmetic practice. *Journal of drugs in dermatology: JDD* 13, 11 (2014), 1312–1314.
- [28] Ziyun Guo. 2023. On the Application of Regional Features in Contemporary Dance Drama: Taking the Dance Drama Shawan Past as an Example. *Journal of Education and Educational Research* 4, 1 (2023), 62–64.
- [29] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*, Vol. 52. Elsevier, 139–183.
- [30] David Hinkamp, Jennie Morton, Donna H Krasnow, Mary Virginia Wilmerding, William J Dawson, Michael G Stewart, Herbert Steven Sims, Jan Potter Reed, Katherine Duvall, and Michael McCann. 2017. Occupational health and the performing arts. *Journal of Occupational and Environmental Medicine* 59, 9 (2017), 843–858.
- [31] Kasper Hornbæk. 2006. Current practice in measuring usability: Challenges to usability studies and research. *International journal of human-computer studies* 64, 2 (2006), 79–102.
- [32] Wen-Shiung Huang, Bing-Kai Hong, Wen-Huang Cheng, Shih-Wei Sun, and Kai-Lung Hua. 2018. A cloud-based intelligent skin and scalp analysis system. In *2018 IEEE Visual Communications and Image Processing (VCIP)*. IEEE, 1–5.
- [33] Cindy Hsin-Liu Kao, Bichlien Nguyen, Asta Roseway, and Michael Dickey. 2017. EarthTones: Chemical Sensing Powders to Detect and Display Environmental Hazards through Color Variation. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) (CHI EA '17). Association for Computing Machinery, New York, NY, USA, 872–883. <https://doi.org/10.1145/3027063.3052754>
- [34] Hsin-Liu Kao, Christian Holz, Asta Roseway, Andres Calvo, and Chris Schmandt. 2016. DuoSkin: rapidly prototyping on-skin user interfaces using skin-friendly materials. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers*. 16–23.
- [35] Hsin-Liu Kao, Manisha Mohan, Chris Schmandt, Joseph A Paradiso, and Katia Vega. 2016. Chromoskin: Towards interactive cosmetics using thermochromic pigments. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 3703–3706.
- [36] A. Keck-Wilhelm, E. Kratz, G. Mildau, M. Ilse, C. Schlee, and D. W. Lachenmeier. 2015. Chemical analysis and risk assessment of prohibited colouring agents in face paint with special regard to CI 15585 (D&C Red No. 9, Pigment Red 53:1). *International Journal of Cosmetic Science* 37, 2 (2015), 187–195. <https://doi.org/10.1111/ics.12181> arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/ics.12181>
- [37] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. 2023. Segment anything. *arXiv preprint arXiv:2304.02643* (2023).
- [38] Laurent Kohl, Arlette Blondeel, and Micheline Song. 2002. Allergic contact dermatitis from cosmetics. *Dermatology* 204, 4 (2002), 334–337.
- [39] CS Lai. 1972. Lead poisoning as an occupational hazard in Chinese opera actors—a case report. *Singapore medical journal* 13, 2 (1972).
- [40] Chengxu Li, Wenmin Fei, Yang Han, Xiaoli Ning, Ziyi Wang, Keke Li, Ke Xue, Jingkai Xu, Ruixing Yu, Rusong Meng, et al. 2021. Construction of an artificial intelligence system in dermatology: Effectiveness and consideration of Chinese Skin Image Database (CSID). *Intelligent Medicine* 1, 02 (2021), 56–60.
- [41] Franklin Mingzhe Li, Franchesca Spektor, Meng Xia, Mina Huh, Peter Cederberg, Yuqi Gong, Kristen Shinohara, and Patrick Carrington. 2022. "It Feels Like Taking a Gamble": Exploring Perceptions, Practices, and Challenges of Using Makeup and Cosmetics for People with Visual Impairments. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 266, 15 pages. <https://doi.org/10.1145/3491102.3517490>
- [42] Tingting Li, Ruihe Qian, Chao Dong, Si Liu, Qiong Yan, Wenwu Zhu, and Liang Lin. 2018. BeautyGAN: Instance-Level Facial Makeup Transfer with Deep Generative Adversarial Network. In *Proceedings of the 26th ACM International Conference on Multimedia* (Seoul, Republic of Korea) (MM '18). Association for Computing Machinery, New York, NY, USA, 645–653. <https://doi.org/10.1145/3240508.3240618>
- [43] Hsueh-Fang Liu. 1997. The Art of facial makeup in Chinese opera. (1997).
- [44] Junyu Liu, Jianfeng Ren, Hongliang Sun, and Xudong Jiang. 2023. Face Recognition on Point Cloud with Cgan-Top for Denoising. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE,

- 1–5.
- [45] Lin Liu, Chen Liang, Yuzhou Xue, Tingqiao Chen, Yangmei Chen, Yufan Lan, Jiamei Wen, Xinyi Shao, and Jin Chen. 2023. An Intelligent Diagnostic Model for Melasma Based on Deep Learning and Multimode Image Input. *Dermatology and Therapy* 13, 2 (2023), 569–579.
- [46] Shilong Liu, Zhaoyang Zeng, Tianhe Ren, Feng Li, Hao Zhang, Jie Yang, Chunyuan Li, Jianwei Yang, Hang Su, Jun Zhu, et al. 2023. Grounding dino: Marrying dino with grounded pre-training for open-set object detection. *arXiv preprint arXiv:2303.05499* (2023).
- [47] Camillo Lugaresi and Jiuqiang Tang. 2019. Hadon Nash, Chris Mc-Clanahan, Esha Uboweja, Michael Hays, Fan Zhang, Chuo-Ling Chang, Ming Guang Yong, Juhyun Lee, et al. Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172* 4 (2019), 15.
- [48] Yasuhiro Matsumura and Honnavara N Ananthaswamy. 2004. Toxic effects of ultraviolet radiation on the skin. *Toxicology and applied pharmacology* 195, 3 (2004), 298–308.
- [49] Yoshie MATSUOKA, Kozo YONEDA, Chieko SADAHIRA, Junko KATSUURA, Tetsuya MORIUE, and Yasuo KUBOTA. 2006. Effects of skin care and makeup under instructions from dermatologists on the quality of life of female patients with acne vulgaris. *The Journal of Dermatology* 33, 11 (2006), 745–752. <https://doi.org/10.1111/j.1346-8138.2006.00174.x> arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1346-8138.2006.00174.x>
- [50] Emanuel Maverakis, Yoshinori Miyamura, Michael P Bowen, Genevieve Correa, Yoko Ono, and Heidi Goodarzi. 2010. Light, including ultraviolet. *Journal of autoimmunity* 34, 3 (2010), J247–J257.
- [51] Irmina Maria Michalek, Emma KT Benn, Florentino Luciano Caetano Dos Santos, Sharon Gordon, Chi Wen, and Bian Liu. 2019. A systematic review of global legal regulations on the permissible level of heavy metals in cosmetics with particular emphasis on skin lightening products. *Environmental research* 170 (2019), 187–193.
- [52] K Morabito, NC Shapley, KG Steele, and A Tripathi. 2011. Review of sunscreen and the emergence of non-conventional absorbers and their applications in ultraviolet protection. *International journal of cosmetic science* 33, 5 (2011), 385–390.
- [53] Maki Nakagawa, Koji Tsukada, and Itiro Sii. 2010. Remote Skincare Advice System Using Life Logs. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (Atlanta, Georgia, USA) (CHI EA '10). Association for Computing Machinery, New York, NY, USA, 3811–3816. <https://doi.org/10.1145/1753846.1754061>
- [54] Le Thi Nhu Ngoc, Vinh Van Tran, Ju-Young Moon, Minhe Chae, Duckshin Park, and Young-Chul Lee. 2019. Recent trends of sunscreen cosmetic: An update review. *Cosmetics* 6, 4 (2019), 64.
- [55] Tam V. Nguyen and Luoqi Liu. 2017. Smart Mirror: Intelligent Makeup Recommendation and Synthesis. In *Proceedings of the 25th ACM International Conference on Multimedia* (Mountain View, California, USA) (MM '17). Association for Computing Machinery, New York, NY, USA, 1253–1254. <https://doi.org/10.1145/3123266.3127926>
- [56] Ayano Nishimura and Itiro Sii. 2014. IMake: Eye Makeup Design Generator. In *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology* (Funchal, Portugal) (ACE '14). Association for Computing Machinery, New York, NY, USA, Article 20, 6 pages. <https://doi.org/10.1145/2663806.2663823>
- [57] Ministry of Culture (PRC). 2018. A total of 348 local opera genres nationwide. [https://www.gov.cn/xinwen/2018-01/02/content\\_5252331.htm](https://www.gov.cn/xinwen/2018-01/02/content_5252331.htm). (Accessed on 07/15/2023).
- [58] Julianne S Oktay. 2012. *Grounded theory*. Pocket Guide to Social Work Re.
- [59] Michelle E Park and Jonathan H Zippin. 2014. Allergic contact dermatitis to cosmetics. *Dermatologic clinics* 32, 1 (2014), 1–11.
- [60] Thies Pfeiffer and Cem Memili. 2016. Model-Based Real-Time Visualization of Realistic Three-Dimensional Heat Maps for Mobile Eye Tracking and Eye Tracking in Virtual Reality. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications* (Charleston, South Carolina) (ETRA '16). Association for Computing Machinery, New York, NY, USA, 95–102. <https://doi.org/10.1145/2857491.2857541>
- [61] Jo Riley. 1997. *Chinese theatre and the actor in performance*. Vol. 3. Cambridge University Press.
- [62] Neelam Sharma, Sumeet Patiyal, Anjali Dhall, Naorem Leimarembi Devi, and Gajendra PS Raghava. 2021. ChAlPred: A web server for prediction of allergenicity of chemical compounds. *Computers in Biology and Medicine* 136 (2021), 104746.
- [63] Chuhan Shi, Zhihan Jiang, Xiaojuan Ma, and Qiong Luo. 2022. A Personalized Visual Aid for Selections of Appearance Building Products with Long-term Effects. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [64] Vimal K Shrivastava, Narendra D Londhe, Rajendra S Sonawane, and Jasjit S Suri. 2016. Computer-aided diagnosis of psoriasis skin images with HOS, texture and color features: a first comparative study of its kind. *Computer methods and programs in biomedicine* 126 (2016), 98–109.
- [65] Sagar Sunkle, Deepak Jain, Krati Saxena, Ashwini Patil, Tushita Singh, Beena Rai, and Vinay Kulkarni. 2021. Integrated “generate, make, and test” for formulated products using knowledge graphs. *Data Intelligence* 3, 3 (2021), 340–375.
- [66] Shuo Tian, Wenbo Yang, Jehane Michael Le Grange, Peng Wang, Wei Huang, and Zhewei Ye. 2019. Smart healthcare: making medical care more intelligent. *Global Health Journal* 3, 3 (2019), 62–65.
- [67] Bantita Trepong, Hironori Mitake, and Shoichi Hasegawa. 2018. Makeup Creativity Enhancement with an Augmented Reality Face Makeup System. *Comput. Entertain.* 16, 4, Article 6 (nov 2018), 17 pages. <https://doi.org/10.1145/3277452>
- [68] Anh Truong, Peggy Chi, David Salesin, Irfan Essa, and Maneesh Agrawala. 2021. Automatic Generation of Two-Level Hierarchical Tutorials from Instructional Makeup Videos. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 108, 16 pages. <https://doi.org/10.1145/3411764.3445721>
- [69] Jao Hsun Tseng and Po Hsien Lin. 2013. A Comparison of the Use of Painted Faces in the Peking Opera with the Use of Masks in Western Theater. In *2013 International Conference on the Modern Development of Humanities and Social Science*. Atlantis Press, 304–308.
- [70] Norimichi Tsumura, Nobutoshi Ojima, Kayoko Sato, Mitsuhiro Shirashi, Hideto Shimizu, Hirohide Nabeshima, Syuuichi Akazaki, Kimihiko Hori, and Yoichi Miyake. 2003. Image-Based Skin Color and Texture Analysis/Synthesis by Extracting Hemoglobin and Melanin Information in the Skin. *ACM Trans. Graph.* 22, 3 (jul 2003), 770–779. <https://doi.org/10.1145/882262.882344>
- [71] UNESCO. 2023. Browse the Lists of Intangible Cultural Heritage and the Register of good safeguarding practices - intangible heritage - Culture Sector - UNESCO. <https://ich.unesco.org/en/lists>. (Accessed on 07/15/2023).
- [72] A Vakunov and D Lagun. 2020. MediaPipe Iris: Real-time Iris Tracking & Depth Estimation. (2020).
- [73] Pat Vatiwutitpong, Sirawich Vachmanus, Thanapon Noraset, and Suppawong Tuarob. 2023. Artificial Intelligence in Cosmetic Dermatology: A Systematic Literature Review. *IEEE Access* 11 (2023), 71407–71425. <https://doi.org/10.1109/ACCESS.2023.3295001>
- [74] Bin Wang, Yu Su, Liyan Tian, Shuchuan Peng, and Rong Ji. 2020. Heavy metals in face paints: Assessment of the health risks to Chinese opera actors. *Science of The Total Environment* 724 (2020), 138163.
- [75] Bin Wang, Liyan Tian, Lili Tian, Xisheng Wang, Yujie He, and Rong Ji. 2023. Insights into Health Risks of Face Paint Application to Opera Performers: The Release of Heavy Metals and Stage-Light-Induced Production of Reactive Oxygen Species. *Environmental Science & Technology* 57, 9 (2023), 3703–3712.
- [76] Bin Wang, Zhaohao Zheng, Ling Chen, Wenhui Zhang, Yujie He, Bing Wu, and Rong Ji. 2023. Transcriptomics reveals key regulatory pathways and genes associated with skin diseases induced by face paint usage. *Science of The Total Environment* 890 (2023), 164374. <https://doi.org/10.1016/j.scitotenv.2023.164374>
- [77] Ding Wang, Jinsheng Kang, Sheng-Feng Qin, and Johannes Birringer. 2016. Cultural-based visual expression: emotional analysis of human face via Peking Opera Painted Faces (POPF). *Multimedia Tools and Applications* 75 (2016), 11865–11891.
- [78] Dakuo Wang, Liuping Wang, Zhan Zhang, Ding Wang, Haiyi Zhu, Yvonne Gao, Xiangmin Fan, and Feng Tian. 2021. “Brilliant AI Doctor” in Rural Clinics: Challenges in AI-Powered Clinical Decision Support System Deployment. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 697, 18 pages. <https://doi.org/10.1145/3411764.3445432>
- [79] Jiahao Wang, Zeyu Xiong, Yicun Duan, Junyu Liu, Zilin Song, and Xiangjun Peng. 2021. The Importance Distribution of Drivers’ Facial Expressions Varies over Time!. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Leeds, United Kingdom) (AutomotiveUI '21 Adjunct). Association for Computing Machinery, New York, NY, USA, 148–151. <https://doi.org/10.1145/3473682.3480283>
- [80] Li Wang, Junying Zhao, Wei Xu, and Renlie Cao. 1999. Investigation of oil-coloured dermatitis of the facial skin. *Journal of Practical Medicine* (1999).
- [81] Tai-Jui Wang, Yu-Ju Lin, and Jun-Liang Chen. 2013. The research on cognition design in Chinese opera mask. In *Cross-Cultural Design. Methods, Practice, and Case Studies: 5th International Conference, CCD 2013, Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013, Proceedings, Part I 5*. Springer, 493–502.
- [82] Yu Wang, Mengru Sun, Yifan Duan, et al. 2021. Metagenomic sequencing analysis for acne using machine learning methods adapted to single or multiple data. *Computational and Mathematical Methods in Medicine* 2021 (2021).
- [83] Huajing Xing. 2019. *Category innovation for facial cleansing/makeup remover wipes using a holistic sensory and consumer research approach in a business environment*. Ph.D. Dissertation.
- [84] Zeyu Xiong, Shihan Fu, and Mingming Fan. 2023. OperARTistry: An AR-based Interactive Application to Assist the Learning of Chinese Traditional Opera (Xiqu) Makeup. arXiv:2311.11269 [cs.HC]
- [85] Zeyu Xiong, Jiahao Wang, Wangkai Jin, Junyu Liu, Yicun Duan, Zilin Song, and Xiangjun Peng. 2022. Face2Statistics: user-friendly, low-cost and effective alternative to in-vehicle sensors/monitors for drivers. In *International Conference on Human-Computer Interaction*. Springer, 289–308.

- [86] Qixin Yan, Chunle Guo, Jixin Zhao, Yuekun Dai, Chen Change Loy, and Chongyi Li. 2022. BeautyREC: Robust, Efficient, and Content-preserving Makeup Transfer. *arXiv preprint arXiv:2212.05855* (2022).
- [87] Yin Yang, Yiping Ge, Lifang Guo, Qiuju Wu, Lin Peng, Erjia Zhang, Junxiang Xie, Yong Li, and Tong Lin. 2021. Development and validation of two artificial intelligence models for diagnosing benign, pigmented facial skin lesions. *Skin Research and Technology* 27, 1 (2021), 74–79.
- [88] Rachida Zegour, Ahror Belaid, Julien Ognard, and Douraid Ben Salem. 2023. Convolutional neural networks-based method for skin hydration measurements in high resolution MRI. *Biomedical Signal Processing and Control* 81 (2023), 104491.
- [89] Cao Zhan, Liang Zheng, Yao Shengying, and Wang Chen. 2018. Occupational Melanosis in Peking Opera Actors: A Report of Two Cases. *Capital Public Health* 12, 4 (2018), 189–190.
- [90] Xiang Zhang, Teng Zhou, and Ka Ming Ng. 2021. Optimization-based cosmetic formulation: Integration of mechanistic model, surrogate model, and heuristics. *AIChE Journal* 67, 1 (2021), e17064.
- [91] Hao Zhou and Sudhir P Mudur. 2007. 3D scan-based animation techniques for Chinese opera facial expression documentation. *Computers & Graphics* 31, 6 (2007), 788–799.

## A DEMOGRAPHIC DATA OF SEMI-STRUCTURED INTERVIEW PARTICIPANTS

Id	Gender	Age	Years of Xiqu Makeup Experience	Opera Type	Avg Face Paint Wearing Duration (hour)	Avg Face Paint Wearing Frequency
1	M	15-18	6-10	Cantonese Opera	1-2	2-3 times / week
2	F	26-35	1-5	Beijing Opera	1-2	1-2 times / month
3	F	15-18	>10	Beijing Opera	4-6	1-2 times / month
4	M	18-25	1-5	Beijing Opera	2-4	1-2 times / month
5	F	18-25	6-10	Yue Opera	>6	>3 times / week
6	F	18-25	1-5	Kunqu	2-4	once a week
7	F	18-25	1-5	Wuxi Opera	4-6	once a week
8	F	>55	>10	Beijing Opera	1-2	2-3 times / week
9	F	36-45	>10	Cantonese Opera	4-6	once a week
10	F	15-18	6-10	Cantonese Opera	4-6	1-2 times / month
11	M	36-45	>10	Kunqu	>6	>3 times / week
12	M	26-35	6-10	Henan Opera	2-4	1-2 times / week
13	F	>55	>10	Cantonese Opera	4-6	2-3 times / week
14	F	>55	6-10	Cantonese Opera	2-4	1-2 times / week
15	M	>55	1-5	Beijing Opera	2-4	once a week

**Table 3: Basic demographic information of the 15 participants in semi-structured interviews**

## B QUANTITATIVE COMPARISON OF EYEVIS AND DAILY MAKEUP REMOVAL

	$r_p_{baseline}$	$r_p_{EyeVis\_mean}$	$r_b_{baseline}$	$r_b_{EyeVis\_mean}$
P1	22.1%	2.8%	18.4%	3.2%
P2	29.7%	5.4%	26.4%	3.7%
P3	25.5%	3.7%	28.1%	4.6%
P4	17.9%	4.4%	21%	2.5%
P5	27.2%	6.9%	29.1%	6.2%
P6	11.4%	1.6%	8.9%	1.3%
P7	14.2%	3.9%	12.2%	4.1%
P8	22.1%	5.2%	24.7%	6.1%
P9	26%	5.4%	21.4%	4.7%
P10	16.4%	3.2%	13.8%	3.1%
P11	9.9%	1.1%	11.6%	1.2%
P12	15.7%	4.4%	18.2%	2.9%

**Table 4: Ratio data of specified regions of the final taken eye image of each participant’s daily makeup removal process (baseline) and using EyeVis (mean value of 5 trials).  $r_p_{baseline}$ ,  $r_b_{baseline}$ : ratio of pink & black face paint area to the whole baseline image.  $r_p_{EyeVis\_mean}$ ,  $r_b_{EyeVis\_mean}$ : ratio of the mean value of pink & black face paint area to the whole image (using EyeVis).**