

# FaceAtlasAR: Atlas of Facial Acupuncture Points in Augmented Reality

Menghe Zhang · Jürgen P. Schulze · Dong Zhang

Received: date / Accepted: date

**Abstract** Acupuncture is a technique in which practitioners stimulate specific points on the body. Those points, called acupuncture points (or acupoints), anatomically define areas on the skin relative to specific landmarks on the body. However, mapping the acupoints to individuals could be challenging for inexperienced acupuncturists. In this project, we proposed a system to localize and visualize facial acupoints for individuals in an augmented reality (AR) context. This system combines a face alignment model and a hair segmentation model to provide dense reference points for acupoints localization in real-time (60FPS). The localization process takes the proportional bone (B-cun or skeletal) measurement method, which is commonly operated by specialists; however, in the real practice, operators sometimes find it inaccurate due to the skill-related error. With this system, users, even without any skills, can locate the facial acupoints as a part of the self-training or self-treatment process.

**Keywords** Augmented reality · Acupuncture point · Face alignment · Hair segmentation

## 1 Introduction

Acupuncture[1] is a form of alternative medicine and a key component of traditional Chinese medicine (TCM). Based

Menghe Zhang  
Department of Computer Science, University of California, San Diego,  
CA, USA,  
E-mail: mez071@eng.ucsd.edu

Jürgen P. Schulze  
Department of Computer Science, University of California, San Diego,  
CA, USA,  
E-mail: jschulze@ucsd.edu

Dong Zhang  
Qilu University of Technology, Shandong, China,  
E-mail: jnzd156@163.com

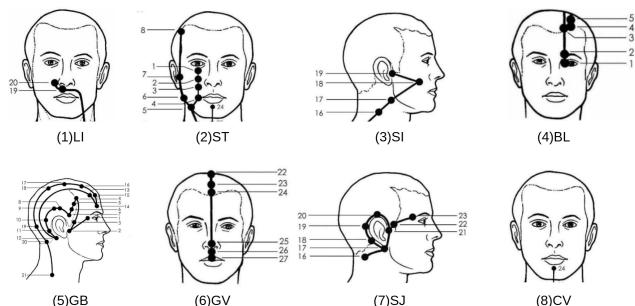


Fig. 1: Facial acupoints grouped by meridians.(1)

on the symptoms, acupuncturists stimulate specific anatomic sites commonly by needling, massaging, or heat therapy. Scientific studies have proved that acupuncture may help ease types of pain that are often chronic such as low-back pain, neck pain, and osteoarthritis/knee pain. The acupoints on the face can help with a variety of conditions both on and off the face, such as jaw tension, headaches, anxiety, and stomach conditions. Figure 1 shows facial acupoints on the face that are categorized by meridians.

Ear reflexology, also known as auricular therapy, is a diagnostic and healing treatment that refers to physical stimulation of the reflex areas of the ears without the use of needles. It also originated from TCM 4000 years ago and has been practiced for centuries. In 1957, Dr Paul Nogier[2] proposed the concept of an inverted fetus map on the external ear, which spread to China and renewed interest in TCM[3]. Later Oleson et al.[4] experimentally evaluate the claims by French and Chinese ear acupuncture that a somatotopic mapping of the body was represented in the external ear. Ear Reflexology is applied for the instant relief of pain[5][6]; weight control[7]; reduction of depression[8] and hypertension. The growing use of auricular acupuncture

for the treatment of chronic pain has occurred in China, Europe, and the United States because of the repeated clinical experience of the effectiveness of this technique. Compared to general body acupuncture, which is more attributed to the increased flow of *qi* along the meridians, this method emphasizes ontogenetic and neurophysiologic connections between auricular reflex points and the central nervous system to explain the somatotopic relations between auricular regions and body pathology. This is based on the facts that within the ear there exists more than 200 nerve innervations and multiple connections with the central nervous system. Every part of the body has an auricular correspondence (similar to the theory of foot reflexology). Therefore, the points for Ear Reflexology are similar to those in ear acupuncture.

However, acupuncture practice relies on experienced acupuncturists to locate the acupoints from body acupuncture maps, which requires long-term theory study and clinical practice. Generally, atlas and models are adopted because its simplicity. However, it loses the effectiveness and lack of intuitive reference. In the ancient time, there exist The Bronze Man[9] as a standard acupuncture training model for students to practice localizing acupoints. Modern professional medical procedure training models improve the reality of practice by adopting skin-like materials. And most recently, with development of virtual reality they created acupuncture training programs that integrate detailed anatomy model and the meridian system[10][11]. On the one hand, the trend of modern training process is to emphasize more anatomical information but lack of a tool to practice on different people; On the other hand, individuals who want to help themselves relieve symptoms with acupoints stimulation but without much medical background, usually find it confusing to localize the targets by natural language description or pictures of a standard model.

To resolve this problem, we designed a system with help of Augmented Reality(AR)[12], in which 3D virtual information are projected into the real world to provide users a more immersive experience. The designated application overlays facial acupoints on the top of the users face, so that users could view and acupressure themselves without any prior knowledge.

Nomenclature and location of auricular points[13] specifies the name and the standard location of the ear acupoints(Fig.2). The goal is to overlay the auricular zone map to users' ears in real time.

Specifically, we employ a deep learning model to annotate 3D landmarks and on a users face together with hair segmentation in real-time to gather reference points for acupuncture points localization. We then align the reference point with acupuncture points defined by proportional bone measurement method. The whole process is implemented via

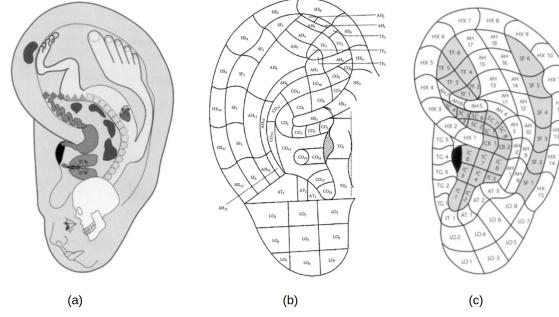


Fig. 2: (a)Somatotopic map originally by Nogier et al.[14]; (b)Schematic diagram of the standard location of auricular points and standard auricle partition zones by Guobiao standards[13]; (c)Auricular zones

MediaPipe, a framework for building cross-platform machine learning solutions. Our system works perfectly on phones with a front camera, without any extra hardware, users could pinpoint and interact with the target acupoints. There are three benefits to this application:

- It takes a conventional localization method while originally defines a scheme to transform the natural language descriptions to mathematical logic expressions.
- It adopts MediaPipe framework to run across platforms in real-time.
- It adapts to different head poses to be robust for users to locate acupoints in different regions.

With FaceAtlasAR, people who have little or no experience in localizing facial acupuncture points can use to relieve their pain or to work on healing their diseases themselves.

## 2 Related Work

### 2.1 Acupuncture point localization

Existing works of acupuncture training applications on AR devices are limited. In 2015, H. Jiang et al.[1] proposed the first acupuncture training application, Acu Glass, on a head-mount display device(HMD) based on Google Glass. They generated the frontal face acupoints based on the height and the width of the input face, plus the distance between the eyes. However, their face landmarks for reference are too limited to adapt to different people and different poses of the face. Other acupoints localization methods like Chen et al.[15][16] fit a 3D Morphable Face Model(3DMM)[17] to a 2D image, and combine facial landmarks and image deformation to estimate acupoints. 3DMM is a powerful tool to build polygonal mesh though, the range of possible predictions is limited by the linear manifold spanned by the PCA basis, which is in turn determined by the diversity of the set

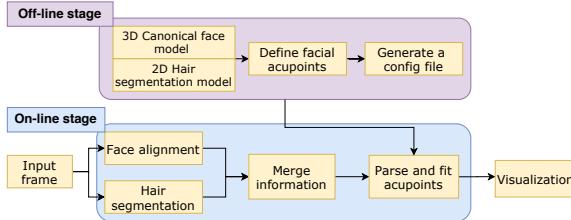


Fig. 3: Workflow of the proposed system

of faces captured for the model[18]. Therefore, manual annotation on a standard 3D model may not correctly fit all kinds of people. Moreover, acupoints are officially defined relative to landmarks, while the deformation process does not guarantee the relativity.

As for the localization methods in practice, Godson and Waridle[19] screened 771 studies and summarized the methods as Directional(F-cun) method, Proportional method, palpation for tenderness, electronic point detectors, and anatomical locations. Usually, more accurate approaches are the next steps of less accurate ones and require extra hardware. For example, one can roughly locate a target by the directional method and then use electronic point detectors to measure the electrical resistance of the skin. There are researches related to this topic and acupoint probing devices available in the market.

## 2.2 Face alignment

Face alignment is a computer vision technology for identifying the geometric structure of human faces in digital images. Bulat and Tzimiropoulos[20] reviewed 2D and 3D face alignment and landmark localization. Existing 2D and 3D datasets annotate a limited set of landmarks. For example, 300-W[21], the most widely-used in-the-wild dataset for 2D alignment, containing LFW[22], HELEN[23], AFW[24], and iBUG[21], annotates only 68 landmarks per face. These landmarks either have distinct semantics of their own or participate in meaningful facial contours. Works based on them are not suitable for our cases. We finally adopted the work of Kartynnik et al.[18], which estimates 3D mesh with 468 vertices in real-time. The vertices are selected manually according to expressive AR effects, thus well suitable for our requirements.

## 3 System Overview

Our facial localization solution utilizes the MediaPipe machine learning pipeline consisting of an off-line stage and an on-line stage(Fig.3).

During the off-line stage:

- **Model selection:** Based on the definition of the referenced points on a face, we choose to blend a pre-trained face alignment model and a pre-trained hair segmentation model. This is because a single model cannot cover the whole set of target regions, while different parts of the regions are in the different reference systems.
- **Acupoints localization:** We firstly localize facial anatomical landmarks, which are designated by the National Standard of the Peoples Republic of China. We then use the proportional bone (B-cun or skeletal) measurement method to locate all acupoints on the face.
- **Data file generation:** The file contains all the information needed for each acupoint/reference point, including name, region, relative location towards a landmark or a reference point. The file is readable for non-technical users, for example, acupuncturists, to correct less accurate acupoint descriptions by natural language.

Then at the on-line stage, the system gets face landmarks together with hair segmentation and merges those results into acupoints generator. The generator gives out the requested acupoints on this frame based on the prior knowledge and then draw on the input face.

We fit the whole process into MediaPipes perception pipeline as a graph of modular components. Each component, called Calculator, solves an individual task like model inference, data transformation, or annotation. We will talk about the implementation details in the next section. We show the graph for our FaceAtlasAR in Figure 4.

The graph consists of two subgraphs: one for face alignment (FaceLandmarkFrontGpu) and the other for hair segmentation (HairSegmentationGpu). From the graph, we see the FlowLimiter guards the whole pipeline process. By connecting the output of the final image to the FlowLimiter with a backward edge, the FlowLimiter keeps track of how many timestamps are currently being processed. The system will down-sample and transform the original image before fusing the input to machine learning models but will draw the results onto the original frames. The next section will show the implementation of each module in detail.

## 4 Facial Acupoints Localizaiton

### 4.1 Face alignment

We adopted a pre-trained TFLite model[18] to infer an approximate 3D mesh representation of a human face.

This process comprises of majorly three steps:

- **Face detection:** The whole frame is first processed by a lightweight face detector to get the face bounding box and several landmarks, thus, to get the rotation matrix of the face. This step only runs until the system finds a face to track or when the system loses tracking.

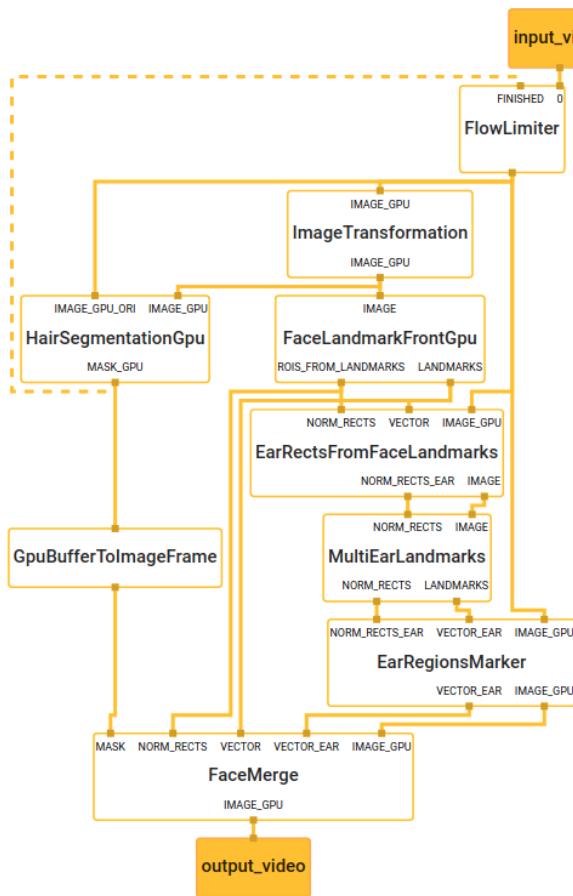


Fig. 4: The graph of FaceAtlasAR

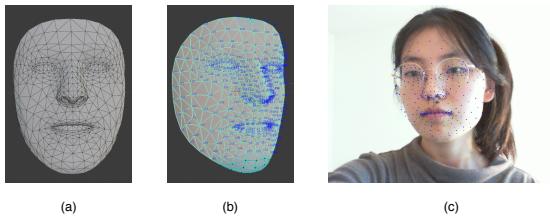


Fig. 5: The generated mesh topology(a), its vertices with index(b), and viewed in AR(c)

- Image transformation:** The image is then cropped by the bounding box and resized to fit into the next step. After this step, the target region is centered and aligned.
- Face landmarks generation:** The pre-trained model produces a vector of 3D landmark coordinates, which subsequently gets mapped back into the original image coordinate system.

Then from a canonical face mesh model, we extract those vertices with semantic meaning as the reference points(Fig.5).

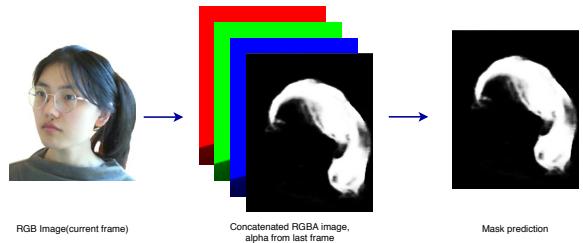


Fig. 6: Hair segmentation module

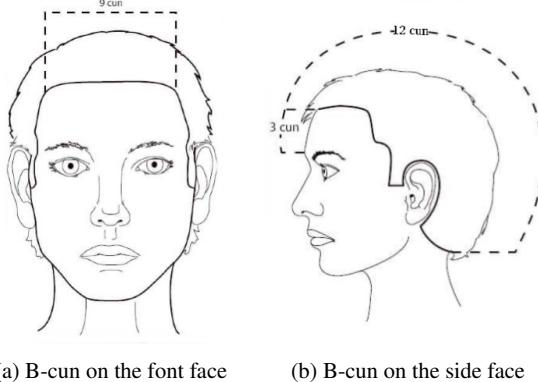


Fig. 7: B-cun on the head from National Standard of the People's Republic of China, Acupoints[26]. Pictures from[27].

#### 4.2 Hair segmentation

Since the center of the frontal hairline is a critical facial anatomical landmark according to the national standard, we adopted a pre-trained model[25] to get the hair segmentation mask. Similar to the face alignment process, the previously generated mask can be fed back to help accelerate the process. Specifically, the mask from the previous round of inference will be embedded as the alpha channel of the current input image(Fig.6).

#### 4.3 Facial acupoints definition

Given the hair mask together with face landmarks, we now could locate facial acupoints based on the B-cun method. This refers to the method of measuring the length and width of each part of the body with the body surface condyles as the main landmark and determining the position of acupoints. Then, a unit cun is the length between the set two bone nodes divided into certain equal parts as the basis for setting acupoints. The facial acupoints bank on the unit cuns definition of the head as shown in Fig.7. To start with, we locate three facial anatomical landmarks in consonance with the national standard. In order to differentiate these three

Channel Name	ID	NameE	Region
ST	2	Sibai	eye
FaceMeshX	FaceMeshY	IsSymmetry	Comments
GetX(RHD3)	GetY(ST1)+0.5*U	TRUE	-

Table 1: Information of an example acupoint Sibai(ST2) in the data file.

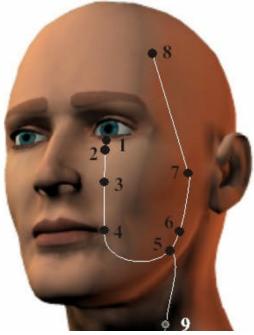


Fig. 8: Illustration of ST channel on the head. Picture from [28]

points to acupoints, we group them in the channel named RHD:

- **RHD1,Yintang:** The midway between the medial ends of the eyebrows.
- **RHD2,Middle of anterior hairline:** Intersection of anterior hairline and anterior midline.
- **RHD3,Pupils**

From Fig.7b, the distance from Yintang to Middle of anterior hairline,  $d_{RHD1 \rightarrow RHD2}$  decides the unit cun as  $uc = d_{RHD1 \rightarrow RHD2}/3$ . On that occasion, like RHD points definition, we could locate all facial acupoints. Table1 shows an example of what information we keep for each acupoint. All points information finally makes up to a data file.

There is one more aspect we want to specify, the channel name, which refers to a unique meridian channel. The meridian system is a concept in TCM about a path through which the life-energy is known as "qi" flows. There are 12 standard and 8 extraordinary meridians, while acupoints are the chosen sites on the meridian system. So we group the acupoints on the same meridian channel and connect them by the flow. For example, the previously stated acupoint, Sibai, belongs to the ST(Foot's Yang Supreme Stomach Meridian) channel. Acupoints on the ST channel and their flow are illustrated in the Fig.8.

## 5 Ear Reflexology Visualization

We started from preparing an annotated template reflexology mapping ear image as  $I$  and its corresponding set of

landmarks  $L_I$ . Then for each frame, the task is divided into several steps and illustrate them each in a paragraph:

- Get cropped ear images(Sec.5.1), each of which as  $J$ , and its landmarks set  $L_J$ (Sec.5.2).
- Make  $J$  the same size as  $I$ , while transform  $L_J$  to  $L_I$  coordinate accordingly. Then  $L_J$  and  $L_I$  become correspondence points.
- Find location of the landmarks in the morphed image.
- Obtain the triangle list  $T$  from  $L_J$  using Delaunay Triangulation method(Sec.5.3).
- Warp and alpha blend image  $I$  and image  $J$  to get the result  $M$ (Sec.5.4).

### 5.1 Ear Detection from Facial landmarks

The measuring of ear position is to find the location of the superior attachment of the pinna. Given facial landmarks in 2D and 3D, the first step is to estimate the ear position and rotation coarsely. In photograph, term 3/4 portrait commonly refers to a shot where 3/4 of the model is visible in the frame. Practically, the model's head is set to be slightly turned away from the camera until the ear opposite the camera is just out of shot. Even though different face and ear shapes make the angles of heads varied, the facing angle is about 45 deg so that the other ear is out of the view. In a side view, it's tilted in relation to the plane of the face. Therefore, we first estimate the head pose from a selected set of face landmarks.

The pose estimation problem refers to estimate its relative orientation and position with respect to a camera and is usually described as Perspective-n-Point(PnP) problem. The goal of such problem is to find the pose of an object with a calibrated camera, and known of  $N$  3D object points and the corresponding 2D projection points in the image. Then we use OpenCV `solvePnP(p3d, p2d, cM, dist)` with selected 2D and 3D landmarks as input. Kartynnik et al.[18] presents a canonical 3D facial mesh where they estimates 3D positions for vertices. Since  $N \geq 4$ , we choose 6 vertices accordingly and illustrate them on the provided canonical mesh(Fig.9). Those points are:

- $P_1$ : Tip of the nose (.0,.0,0)
- $P_2$ : Middle point of the chin (.0, -381.915, 129.21)
- $P_3$ : Left corner of the left eye (-222.293, 189.5445, 215.109)
- $P_4$ : Right corner of the right eye (222.293, 189.5445, 215.109)
- $P_5$ : Left corner of the mouth (-122.8105, -160.786, 159.586)
- $P_6$ : Right corner of the mouth (122.8105, -160.786, 159.586)

Then we get a rotation vector  $\mathbf{V}(x,y,z)$  indicating the head pose towards the camera and the rules to generate ear bounding boxes are:

- If the rotation angle on Y-axis is within 45°, both ears might be seen. Otherwise, only one side of the ears can be seen.

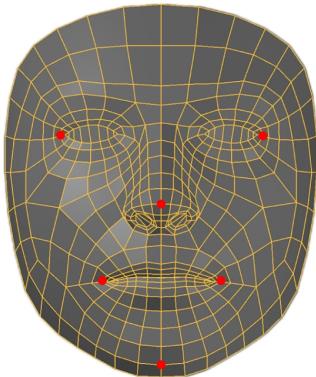


Fig. 9: The predicted mesh topology given by[18] annotated with the selected 6 points.

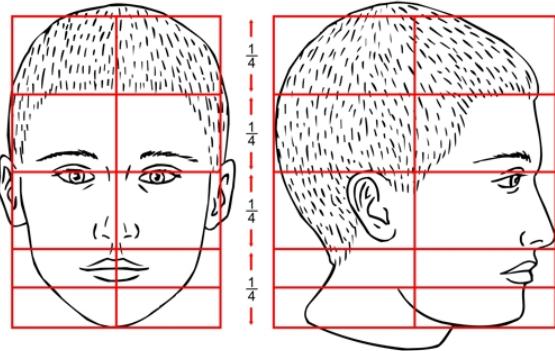


Fig. 10: Proportion of the head[29]

- Rotation on X-axis determines the vertical range while rotation on Z-axis is used to calculate the horizontal range.

Figure 10 illustrates the proportions of head to be divided into four equal quarters horizontally. The third quarter contains most of the features. At the top of this section the eyes are usually level with the ears, and at the bottom the nose is roughly level with the ear lobes. We picked the middle point between the two eyes, nose tip and two selected landmarks on face border to determine the coarse bounding rect of the ear. Given the current gpu frame and the coarse potential ear rectangles inferring from face alignment, there is an optional FineCropEar module(Fig.13) before actual landmarks localization that crops the regions of interest(ROIs) much precisely and rule out false detections. Figure 11 presents the steps to obtain a more precise ROI from a coarse one. First we convert the image from RGB space to the HSV color space and apply HSV color range limitation to the HSV frame. The lower- and upper-bound of HSV color is given by the color range in a sample patch between ear and face. The patch is picked to decide the skin HSV color boundaries thus to leverage different lighting conditions and ethnicities. Then an opening operation to remove small false-positive

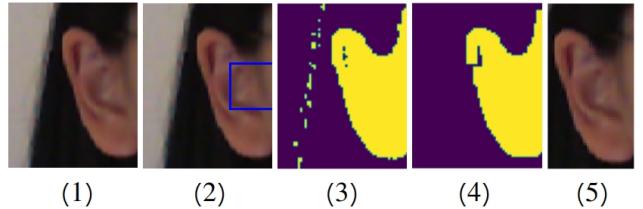
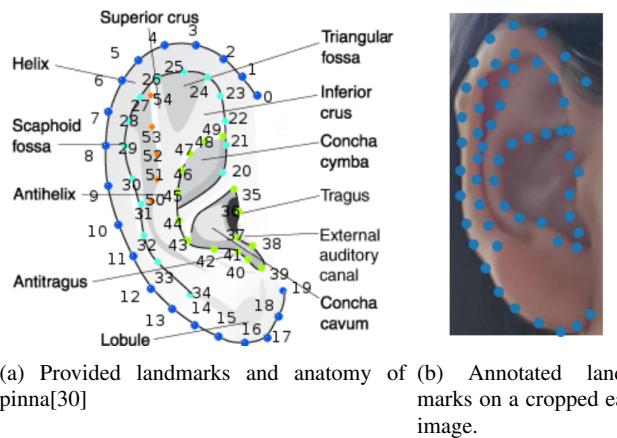


Fig. 11: FineCropEar module process.(1)Input coarse cropped ear image; (2)Image patch to decide skin color boundaries; (3)Ear mask with skin color detection; (4)Mask after an opening operation; (5)Fine cropped ear image.



(a) Provided landmarks and anatomy of pinna[30] (b) Annotated landmarks on a cropped ear image.

Fig. 12: Annotated 55 landmarks on ears from "In-the-wild" ear database[30].

skin regions. The contour with the biggest bounding rectangle is the final target ROI.

## 5.2 Ear Landmarks localization

The inner structure of the human ear is very distinctive and complex. When captured in unconstrained conditions, it is vital that the landmarks dataset is built from deformable ears in-the-wild. "In-the-wild" ear database(Fig.12a) annotates 55 ear landmarks based on the anatomical regions including ascending helix (0- 3), descending helix (4-7), helix (8-13), ear lobe (14-19), ascending inner helix (20-24), descending inner helix (25-28), inner helix (29-34), tragus (35-38), canal (39), antitragus (40-42), concha (43-46), inferio crus (47-49) and supperio crus (50-54). Figure 12 shows a template cropped ear image overlayed with those landmarks

For this task, Hansley et al.[31] provide an unconstrained ear recognition framework that works on ears-in-wild scenarios. The major improvement is to resolve the problem of pose variation by designing a two-stage landmark detector. We modified the original framework with supports to

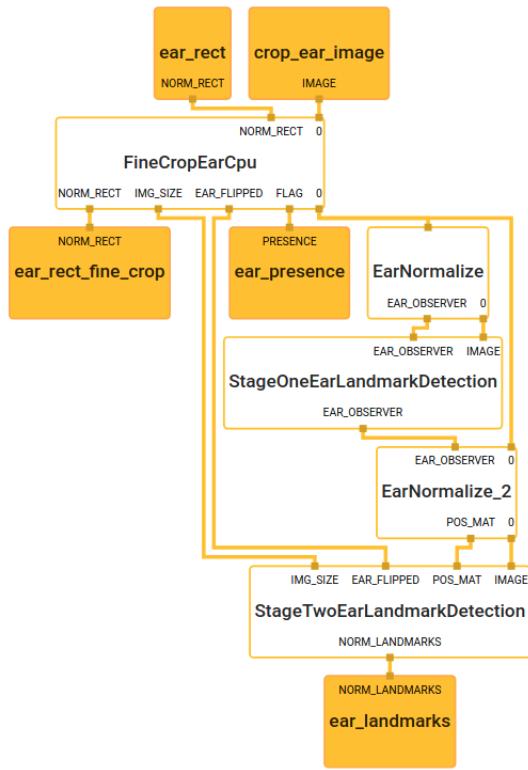


Fig. 13: Single ear landmarks detection Subgraph

project the generated 55 landmarks back to the original image and construct the EarLandmarksSubgraph as shown in Figure 13..

Figure 14 shows a diagram of their framework and an example along with. Compared landmarks results from the first stage(Fig.14(5)) and the second stage(Fig.14(7)), stage one detector is more robust to pose variations while stage two detector is more accurate for well normalized images.

### 5.3 Delaunay Triangulation

Each auricular zone is based upon the principal anatomical regions, in this occasion, anatomical ear landmarks. We first construct triangles from these points and warp to achieve auricular zone mapping. Specifically, we use Delaunay Triangulation[32] that subdivides each input image into a triangular mesh. Given a set of ear landmarks  $P$ , Delaunay triangulation describes the connected triangles' vertices  $DT(P)$  where no point in  $P$  lies within the circumcircle of any other triangle in  $DT(P)$ . The reason to take Delaunay triangulation is because it maximizes the smallest angles and avoids skinny triangles[33].

Then these triangular parts are warped and blended to produce the morph from the annotated template auricular zone

mapping to the detected ear image frames. The triangulation is shown in Figure. Notice the triangles in the two images capture approximately similar regions. Thereby, this step tranforms the point correspondences to region correspondences.

### 5.4 Ear Morphing and Alpha Blending

The triangulation step provides a triangle list  $T$ . With warping operation applied on each of the triangle  $T_i$  in  $T$ , we can blend  $I$  and  $J$  to  $M$  by each triangle. The amount of blending is controlled by the  $\alpha$  parameter where

$$M(x,y) = (1 - \alpha)I(x,y) + \alpha J(x,y) \quad \text{where} \quad 0 \leq \alpha \leq 1(1)$$

The steps go for each of triangle  $T_i$  in  $T$ :

- Calculate the affine transform that maps the  $(T_{i1}, T_{i2}, T_{i3})$  in  $L_I$  and  $L_J$  to that in  $L_M$ .
- Transform all pixels inside the triangle of  $I$  and  $J$  to the morphed image  $M$ .
- Alpha blend the triangle patches and saved to  $M$ .

The results of applying the above technique are shown in Figure 16. The morph image is a 70% blend of the input frame on the template mapping one. We see that since all the landmarks lie on the ear anatomically, the features are well aligned. Thus the triangles and the auricular zones are alinged as a result.

## 6 Results

We achieve real-time performance on both desktop and mobile devices by designing the pipeline properly. We show the final application below and compare the performance on different platforms.

### 6.1 Android Application

Figure 17 presents the screenshots from our FaceAtlasAR android app. Here we show the visualization of acupoints grouped by meridian channels in different poses. Thanks to the robustness of face alignment towards occlusion, users would not find problems pointing or pressing a target acupoint(Fig.18).

### 6.2 Accuracy

We setup an experiment on Samsung S10( $2280 \times 1080$ ) to validate the accuracy of our FaceAtlasAR to localize 4 reference points and 69 acupuncture points on face by comparing the mean pixel errors between ground truth positions and

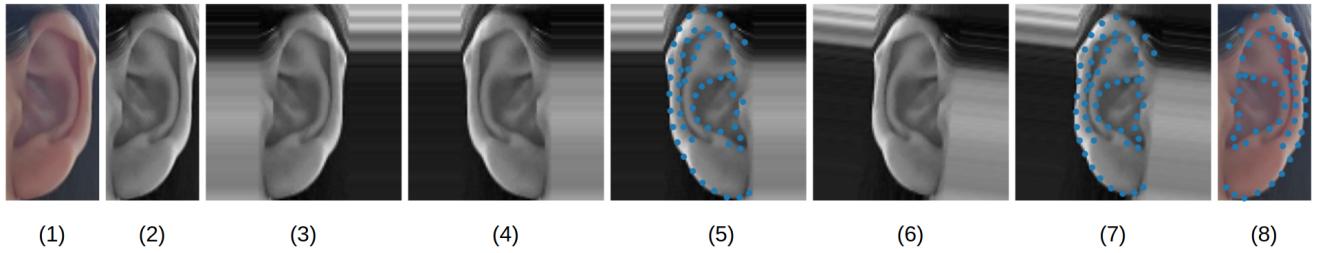


Fig. 14: Diagram and example of our cropped ear landmarks localization based on Hansley et al.[31]. (1)Cropped ear image; (2)Convert the image to grayscle; (3)Coarse parameter estimation and image normalization; (4)Side classification and flip if needed; (5)First stage landmarks detection; (6)Adjustments based on (5) and normalization; (7)Second stage landmarks detection; (8)Backproject landmarks to (1). The original frameworks contains step (1)-(7) while our pipeline drops step (4) thanks to the face alignment step and adds position mapping to support (8).

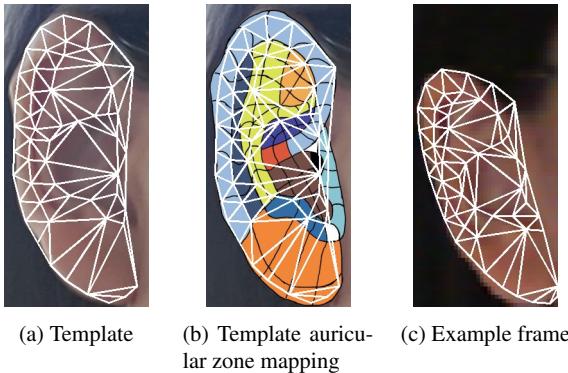


Fig. 15: Triangulations

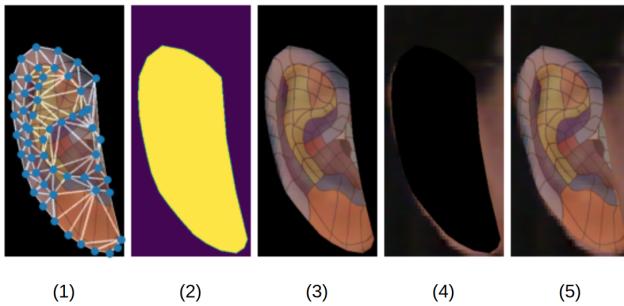


Fig. 16: Ear morph and alpha blending.(1)Morph result  $M$  with re-mapped landmarks and triangles; (2)Mask of  $M$ ; (3)Foreground of the morph image; (4)Background from the input frame; (5)Final blended result.

Table 2: Acupoints categorized by localization times

Method	Direct	One-time proportional	Multiple-times proportional
Reference points	3	1	1
Acupoints	38	16	15



Fig. 17: Android application screenshots for displaying acupoints grouped by meridian system in different poses

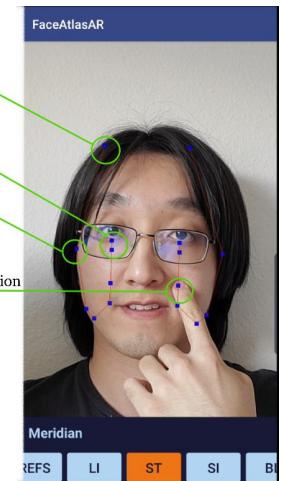


Fig. 18: User presses an acupoint on ST meridian channel

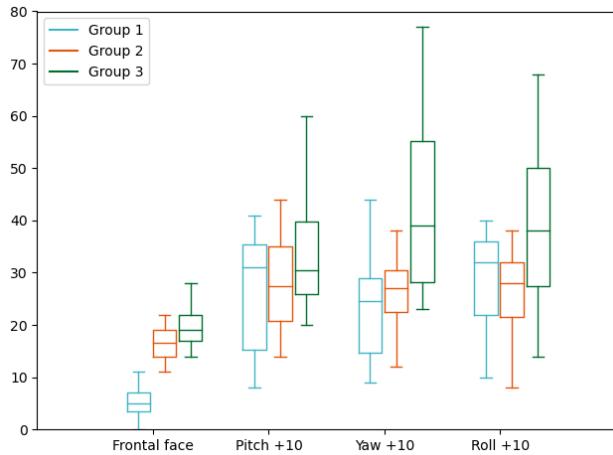


Fig. 19: Mean pixel errors of localization

the localization results. Initially, we assign all the target positions into three groups by the localization complexity as shown in the Table 2.

We then investigate the mean pixel errors in 3 groups. For each localization points, we measure the pixel errors in 4 different poses: frontal face ( $0^\circ$ ), pitch (X-axis+ $10^\circ$ ), roll (Y-axis + $10^\circ$ ), yaw (Z-axis + $10^\circ$ ) to get the mean value. The results are shown in Figure 19.

From the results we found that some of the localizations are less precise than others. Especially, the system performs worse on group 2 than group 1 and group 3 than group 2. Accordingly, multiple times of proportional calculations that rely on the cun system add inaccuracy to the results. The system tolerates to different angles to some extent. However, a target point will be more well localized when it is point straight to the camera without any angles.

### 6.3 Performance

We evaluate the performance of our pipeline in three major components: face alignment, hair segmentation, and acupoints generation. Since the application runs on multiple platforms, we compare its performance on a desktop with Nvidia GeForce RTX 2070 SUPER and on a Samsung S10. The input images for two TFLite models are both in full size  $512 \times 512$ . We see that in this case, Samsung S10 still runs in full frame rate (60 FPS). More detailed comparison is as in Table 3. We also evaluate the performance of data file parsing (Table 3), which only runs once at the setup stage.

## 7 Discussion

From the results we could see that:

Table 3: Data file parsing performance

Device/Time,ms	File parser
Desktop	0.102
Samsung S10	0.279

- Our system can properly display the requested acupoints on selected meridian channels.
- The system tolerates movements very well, while endures tilt and rotation to some degree.
- The benefit from head rotation is that: when the acupoints are hidden in one view, they will be visible and more accurate in another one. For example, the frontal face hides acupoints in the ear region; thus, to view them around the left ear properly, the user needs to turn his/her head so that the left ear faces the camera.

Our next step is to improve the face alignment performance on the side of the head. There are dozens of acupuncture points around the ears that represent specific domain and functions of the body. However, most face alignment jobs only require a small set of landmarks. Even though the model we adopted can estimate 3D mesh with 468 vertices, it still neglects both sides of the head and loses some accuracy at the cheeks and chins. Therefore, we could only roughly estimate those acupoints positions based on the proportional relations to other landmarks on face, which is less meticulous.

## 8 Conclusion

In this paper, we proposed FaceAtlasAR, an end-to-end facial acupoints tracking solution that achieves real-time performance on mobile devices. Our pipeline integrates a face alignment model with a hair segmentation model. The high accuracy of the estimation and the robustness of the system empower users with little experience in acupuncture to interact with facial acupoints. Future work comes to improving the accuracy even further in the ear region since the face alignment only gives a little information towards the face edge near the ear region. Also, 3D interaction should be considered for users to gain a more immersive experience. For example, we could track users hands/fingers while they are interacting with a target acupoint.

## References

1. F. Li, T. He, Q. Xu, L.T. Lin, H. Li, Y. Liu, G.X. Shi, C.Z. Liu, *Pain Medicine* **16**(10), 1905 (2015)
2. P. Nogier, *Bulletin de la Société d'Acupuncture* (20), 51 (1956)
3. L. Gori, F. Firenzuoli, *Evidence-Based Complementary and Alternative Medicine* **4**(S1), 13 (2007)

Table 4: Application performance on a desktop and a mobile device

Device/Time.ms	Hair segmentation inference	Face alignment inference	Generation	Overall Time
Desktop	1.188	3.376	0.135	10.56
Samsung S10	50.279	14.673	0.512	84.758

4. T.D. Oleson, R.J. Kroening, D.E. Bresler, *Pain* **8**(2), 217 (1980)
5. T. Oleson, W. Flocco, *Obstetrics and Gynecology* **82**, 906 (1993)
6. M.I. Weintraub, R. Mamani, M.S. Micozzi, *Complementary and integrative medicine in pain management* (Springer Publishing Company, 2008)
7. J. Yao, L. Chen, L. Zhang, S. Zhou, Q. Zheng, X. Feng, X. You, L. Zhang, Y. Li, *Medicine* **98**(34) (2019)
8. C. Wu, P. Liu, H. Fu, W. Chen, S. Cui, L. Lu, C. Tang, *Medicine* **97**(52) (2018)
9. A. White, E. Ernst, *Rheumatology* **43**(5), 662 (2004)
10. R. Kanehira, A. Shoda, M. Yagihashi, H. Narita, H. Fujimoto, in *2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery*, vol. 4 (IEEE, 2008), vol. 4, pp. 665–668
11. E. Yau. Virtual reality helps chinese medicine students learn acupuncture and doctors treat cancer. URL <https://www.scmp.com/lifestyle/health-wellness/article/2169092/virtual-reality-helps-chinese-medicine-students-learn>
12. R.T. Azuma, *Presence: teleoperators & virtual environments* **6**(4), 355 (1997)
13. Nomenclature and location of auricular points. Standard, Chinese Acupuncture Society, China (2008)
14. P.M. Nogier, A. Cousino, M. Graff, *From auriculotherapy to auriculomedicine* (Maisonneuve, 1983)
15. Y.Z. Chen, C. Maigre, M.C. Hu, K.C. Lan, in *Proceedings of the 8th ACM on Multimedia Systems Conference* (2017), pp. 239–241
16. K.C. Lan, G.S. Li, J.X. Zhang, in *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems* (2018), pp. 384–385
17. P. Huber, G. Hu, R. Tena, P. Mortazavian, P. Koppen, W.J. Christmas, M. Ratsch, J. Kittler, in *Proceedings of the 11th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* (2016)
18. Y. Kartynnik, A. Ablavatski, I. Grishchenko, M. Grundmann, arXiv preprint arXiv:1907.06724 (2019)
19. D.R. Godson, J.L. Wardle, *Journal of acupuncture and meridian studies* **12**(2), 52 (2019)
20. A. Bulat, G. Tzimiropoulos, in *Proceedings of the IEEE International Conference on Computer Vision* (2017), pp. 1021–1030
21. C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, M. Pantic, in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (2013), pp. 896–903
22. P.N. Belhumeur, D.W. Jacobs, D.J. Kriegman, N. Kumar, *IEEE transactions on pattern analysis and machine intelligence* **35**(12), 2930 (2013)
23. V. Le, J. Brandt, Z. Lin, L. Bourdev, T.S. Huang, in *European conference on computer vision* (Springer, 2012), pp. 679–692
24. X. Zhu, D. Ramanan, in *2012 IEEE conference on computer vision and pattern recognition* (IEEE, 2012), pp. 2879–2886
25. A. Tkachenka, G. Karpiak, A. Vakunov, Y. Kartynnik, A. Ablavatski, V. Bazarevsky, S. Pisarchyk, arXiv preprint arXiv:1907.06740 (2019)
26. Unknown. National standard of the people's republic of china, acupoints (2005). <https://musculoskeletalkey.com/ynsa-basic-points/>
27. Unknown. Ynsa basic points (2017). <https://musculoskeletalkey.com/ynsa-basic-points/>
28. Unknown. Atlas of acupuncture points (2007). [www.AcupunctureProducts.com](http://www.AcupunctureProducts.com)
29. Artyfactory. The proportions of the head. URL <https://www.artfactory.com/portraits/pencil-portraits/proportions-of-a-head.html>
30. Y. Zhou, S. Zaferiou, in *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)* (IEEE, 2017), pp. 626–633
31. E.E. Hansley, M.P. Segundo, S. Sarkar, *IET Biometrics* **7**(3), 215 (2018)
32. D.T. Lee, B.J. Schachter, *International Journal of Computer & Information Sciences* **9**(3), 219 (1980)
33. D.M. Mount, Lecture Notes, University of Maryland pp. 1–122 (2002)