

# FaceAtlasAR: Atlas of Facial Acupuncture Points in Augmented Reality

Menghe Zhang<sup>1</sup>, Jürgen P. Schulze<sup>1</sup> and Dong Zhang<sup>2</sup>

<sup>1</sup>Computer Science and Engineering, University of California, San Diego, 9500 Gilman Dr, La Jolla, 92093, CA, USA.

<sup>2</sup>Automation, Qilu University of Technology, 3501 Daxue Rd, Jinan, 250353, Shandong, China.

Contributing authors: [mez071@eng.ucsd.edu](mailto:mez071@eng.ucsd.edu); [jschulze@ucsd.edu](mailto:jschulze@ucsd.edu); [jnzd156@163.com](mailto:jnzd156@163.com);

## Abstract

Acupuncture is a technique in which practitioners stimulate specific points on the body. These points, called acupuncture points (or acupoints), anatomically define areas on the skin relative to some landmarks on the body. Traditional acupuncture treatment relies on experienced acupuncturists for precise positioning of acupoints. A novice typically finds it difficult because of the lack of visual cues. This project presents FaceAtlasAR, a prototype system that localizes and visualizes facial acupoints in an augmented reality (AR) context. The system aims to 1) localize facial acupoints and auricular zone map in an anatomical yet feasible way, 2) overlay the requested acupoints by category in AR, and 3) show auricular zone map on the ears. We adopt Mediapipe, a cross-platform machine learning framework, to build the pipeline that runs on desktop and Android phones. We perform experiments on different benchmarks, including “In-the-wild”, AMI ear datasets, and our own annotated datasets. Results show the localization accuracy of 95% for facial acupoints, 99% / 97% (“In-the-wild” / AMI) for auricular zone map, and high robustness. With this system, users, even not professionals, can position the acupoints quickly for their self-acupressure treatments.

**Keywords:** Augmented Reality, Acupuncture, Mediapipe, Face alignment, Ear Morphology

## 1 Introduction

Acupuncture[1] is a key component of traditional Chinese medicine (TCM). Based on the symptoms, acupuncturists stimulate specific anatomic sites commonly by needling, massaging, or heat therapy. Those anatomic sites, called acupuncture points (or acupoints), lie on different meridians, where the life energy (called “qi”) flows. Recent research presents the meridians by capturing a fluorescein migratory path proximately along the meridian[2]. Scientific studies have confirmed the effectiveness of acupuncture for many conditions[3][4][5]. Acupuncture practice relies on experienced acupuncturists to localize the acupoints from body acupuncture maps, which

requires long-term theory study and clinical practice. However, patients could do needless-acupuncture treatments like acupressure and cupping as long as they can quickly position the acupoints themselves. Generally, there are standardized atlas and models. Yet, a novice usually finds it confusing to localize the targets by natural language description or pictures of a standard model.

Recently, Augmented Reality (AR)[6] is considered a great way to solve this problem. By projecting the acupoints onto the skin, those AR applications provide users with more intuitive guidance. For facial acupuncture, H. Jiang et al.[7] proposed the first acupuncture training application, Acu Glass, on a

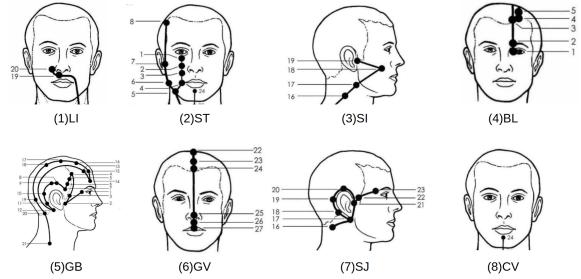
head-mount display device (HMD) based on Google Glass. They generated the frontal face acupoints based on the height and the width of the input face, plus the distance between the eyes. Thus, the face landmarks for reference are too limited to adapt to different people and different poses. Another widely accepted solution is fitting a 3D Morphable Face Model (3DMM)[8] to a 2D image, like Chen et al.[9][10]. They combine facial landmarks and image deformation to estimate acupoints. 3DMM is a powerful tool to build polygonal mesh though, manual annotations on a standard 3D model may not correctly fit all kinds of faces[11]. Moreover, acupoints are officially defined relative to anatomical landmarks, while the deformation process does not guarantee the relativity.

Besides facial acupuncture, ear acupuncture, also known as auricular acupuncture, has also been applied for a long time. It is a micro-acupuncture technique similar to ear reflexology, which is also a diagnostic and healing treatment that refers to physical stimulation of the reflex areas of the ears without the use of needless. We incorporate ear acupuncture guidance in our application with the auricular zone map overlaid on users' ears in AR to help them find the target reflex points. As far as we researched, there has been no report about visualizing auricular zone maps in real-time.

Our previous work proposed FaceAtlasAR, an augmented reality application that projects acupoints on the face in AR. In this paper, we extend it to show auricular zone maps in the meantime. And finally, extensive experiments are conducted to show its effectiveness and robustness.

Our system pipeline is built upon MediaPipe[12], an open-source framework for cross-platform machine learning solutions. For facial acupoints localization, we fuse face alignment and hair segmentation models into the framework pipeline. The resulting masks and face landmarks are applied to calculate reference points and acupoints in an anatomical yet feasible way. Meanwhile, another two deep learning models gather anatomical ear landmarks for auricular zone mapping. Our system works perfectly on Android phones with front cameras. Without any extra hardware, users who have little or no experience practicing acupuncture could relieve the symptoms by doing needles self-treatments.

Compared to other AR acupoints localization projects, our work develops an innovative solution because:



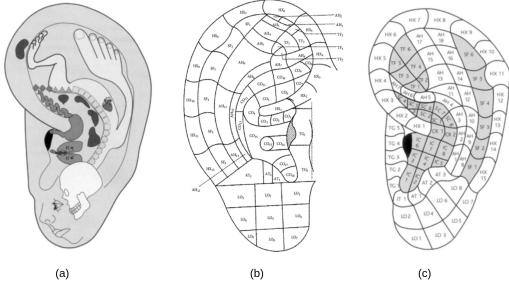
**Fig. 1:** Facial acupoints grouped by meridians[13]. (1) Large Intestine Channel (LI), (2) Stomach Channel (ST), (3) Small Intestine Channel (SI), (4) Bladder Channel (BL), (5) Gall Bladder Channel (GB), (6) The Governor Vessel Channel (GV), (7) Sanjiao Channel (SJ), and (8) Conception Vessel Channel (CV).

- It takes a standard localization method while creatively defining a scheme to transform the natural language descriptions to mathematical logic expressions.
- It is the first AR application that extends the facial acupuncture map to the auricular zone map.
- It adopts MediaPipe as an end-to-end machine learning framework to run across platforms in real-time.

## 2 Background

### 2.1 Facial and Ear Acupuncture

Facial acupuncture can help with various conditions, such as jaw tension, headaches, anxiety, and stomach conditions. The selection of facial acupoints is based on Chinese National Standard (CNS) *Nomenclature and Location of Acupuncture Points* (GB/T 12346-2006) and WHO *Standard Acupuncture Point Locations in the Western Pacific Region*. Figure 1 visualises facial acupoints categorized by meridians. Ear acupuncture (ear reflexology or auricular therapy), originated from TCM 4000 years ago, has been practiced for centuries. In 1957, Dr. Paul Nogier[14] proposed the concept of an inverted fetus map on the external ear (Fig. 2 (a)), which spread to China and renewed interest in TCM[15]. Later Oleson et al.[16] experimentally evaluated the claims by French and Chinese ear acupuncture. He then proposed a somatotopic mapping of the body in the external ear. Nowadays, ear acupuncture is applied for the instant relief of pain[17][18], weight control[19], reduction of depression[20], and hypertension. The growing use



**Fig. 2:** (a) Somatotopic map originally from Nogier et al.[22], (b) Auricle zone system by CNS GB/T 13734-2008[21], (c) Revisions of the auricular nomenclature system developed by Oleson based upon the recommendations of the 1990 WHO nomenclature committee.

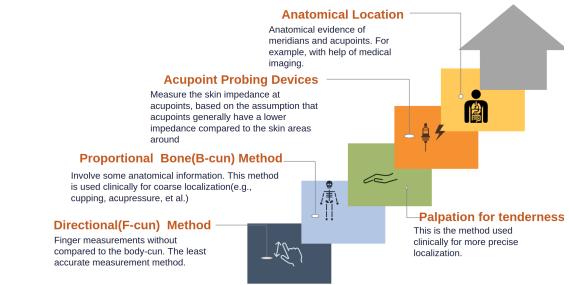
of ear acupuncture for chronic pain has occurred in China, Europe, and the United States because repeated clinical experience has demonstrated the effectiveness of this technique. Ear acupuncture emphasizes neurophysiologic connections between auricular reflex points and the central nervous system. There exist more than 200 nerve innervations in an ear and multiple connections with the central nervous system. Every part of the body has an auricular correspondence, which explains the somatotopic relations between ear regions and body pathology.

Based upon the recommendations of the 1990 WHO auricular nomenclature committee, Figure 2 (b) presents auricular zone system from CNS Nomenclature and location of auricular points (GB/T 13734-2008)[21] while Figure 2 (c) shows the revised UCLA nomenclature system.

## 2.2 AR Assisted Training Solutions

In ancient times, The Bronze Man[23] was a standard acupuncture training model for students to practice localizing acupoints. Modern professional medical procedure training models enhanced the realism by adopting skin-like materials. Most recently, with the development of virtual reality, they created acupuncture training programs that integrate detailed anatomy models and the meridian system[24][25]. The trend of the latest training process is to emphasize more anatomical information. Yet, most of them lack a tool to practice on different bodies.

To achieve that, one needs to focus on: (1) a feasible yet accurate acupoint localization method, (2) a

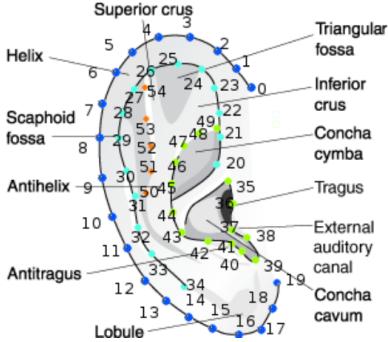


**Fig. 3:** Summary of acupoint localization methods.

fast and robust face alignment tool, and (3) an ear anatomical landmark detection method.

**Acupoints localization** Godson and Wardle[26] screened 771 studies and summarized the methods as Directional (F-cun) method, Proportional Bone (B-cun) method, Palpation for tenderness, using electronic point detectors, and applying anatomical locations. Studies focusing on various localization methods are in Figure 3. All the methods rely at least to some extent on the correct identification of anatomical landmarks. Though the acupoints located in proximity to anatomical landmarks are less variant[27], as the staircase forward, the precision increases while the feasibility drops. Usually, more accurate approaches are the next steps of less accurate ones and require extra hardware. For example, one can find a rough position with the F-cun method, and then uses an electronic point detector to make it more accurate. Shortly, different approaches are just suitable for different scenarios. Generally, there is a balance between the accuracy and the requirements of extra hardware to localize acupoints. For needle acupuncture operations, precision comes first because misplacement can cause serious issues. Inversely, feasibility is the most important for candidates who want to do needless acupressure immediately. Among the five localization methods, some existing AR acupoint localization applications[28] employ the F-cun method (stair 1) in favor of the feasibility. However, the B-cun method is the basis of current standardized acupoints definitions. Thus, in clinics, acupuncturists usually prefer the B-cun method (stair 2) together with the Palpation for tenderness (stair 3). Professional training processes[24][25] also prefer this method because it involves certain anatomical information.

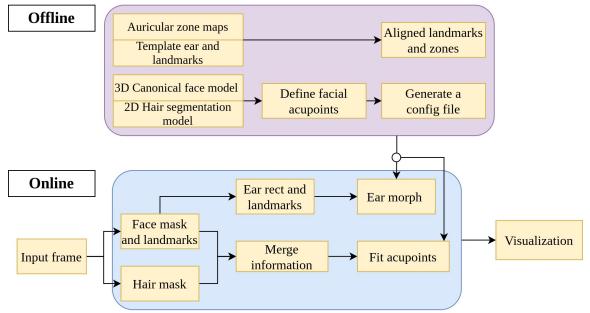
**Face alignment** is a computer vision technology for identifying the geometric structure of human faces in digital images. Bulat and Tzimiropoulos[29] reviewed



**Fig. 4:** Provided landmarks and anatomy of pinna from “In-the-wild” dataset[40]

2D and 3D face alignment and landmark localization. Existing 2D and 3D datasets annotate a limited set of landmarks. For example, 300-W[30], the most widely-used in-the-wild dataset for 2D alignment, containing LFW[31], HELEN[32], AFW[33], and iBUG[30], annotates only 68 landmarks per face. These points only localize some semantic regions around the eyes, eyebrows, nose, mouth, chin, and jaw, which are not dense enough for our cases. We finally adopted the work of Kartynnik et al.[11], which estimates 3D mesh with 468 vertices in real-time. The vertices are selected manually according to expressive AR effect experiments, hence well suitable for our requirements.

**Ear landmarks** With the advances in computer vision, recent studies make great progress working on ear images acquired from real-world unconstrained settings[34][35][36]. Existing databases collect unconstrained ears for recognition, including WPUTE[37], UBEAR[38], and IIT[39]. These datasets are collected under a controlled environment and without annotated structures. The inner structure of the human ear is very distinctive and complex. Zhou’s “In-the-wild” ear database[40] firstly builds statistical deformable models with unconstrained ear samples. It annotates 55 ear landmarks based on the anatomical regions (Fig. 4), 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), inferior crus (47-49) and superior crus (50-54).



**Fig. 5:** Workflow of the proposed system.

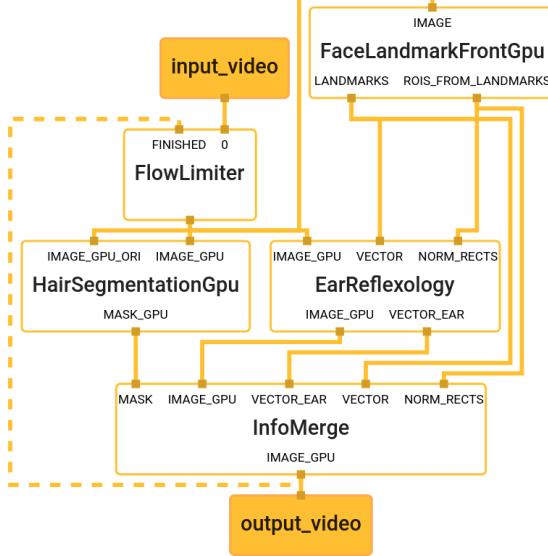
### 3 System Overview

The overall procedure of the proposed system is shown in Figure 5.

It contains an offline phase and an online phase. The offline stage prepares a facial acupoints map and two auricular zone maps, describing acupoints relative to anatomical landmarks. Information of the facial acupoints, which includes name, region, and relative position, is stored in a configuration file. The file is readable for non-technical users, for example, acupuncturists, to correct less accurate acupoint descriptions by natural language. The auricular template also connects zone maps in Figure 2 to the ear landmarks.

During the online phase, the face alignment module generates face landmarks, from which it estimates the coarse ear bounding boxes. Then after the ear morphology step, we see the auricular zone maps fitted on users’ ears. Meanwhile, the pipeline takes the configuration file and hair masks as input and draws the acupoints with prior knowledge.

MediaPipe describes the perception pipeline as a graph of modular components. The smallest component is a Calculator, solving an individual task like model inference, data transformation, or annotation. Reusable calculators group to graphs or sub-graphs, taking charge in a certain task. The main graph is mainly composed of subgraphs and calculators. We show our main graph in Figure 6. To begin with, the FlowLimiter guards the whole pipeline process and keeps track of the timestamps. Then, it consists of three major subgraphs: FaceLandmarks and HairSegmentation subgraphs for Facial Acupoints Localization module (Sec. 4), and EarReflexology for Auricular Zone Visualization module (Sec. 5). We will elaborate on the implementation details in the next sections.

**Fig. 6:** FaceAtlasAR main pipeline graph

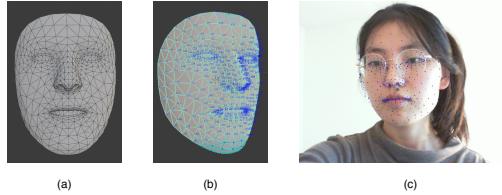
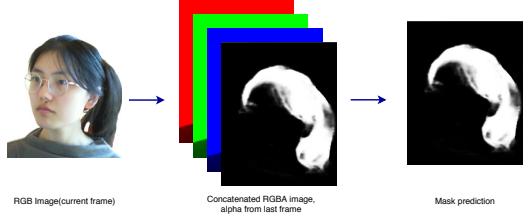
## 4 Facial Acupoints Localizaiton

### 4.1 Face Alignment

We adopted a pre-trained TFLite model[11] to infer an approximate 3D mesh representation of a human face. This process majorly comprises of three steps:

- **Face detection:** A lightweight face detector processes the whole frame to get the face bounding box and several landmarks, determining the rotation matrix of the face. This step only runs until the system finds a face to track or when the system loses track.
- **Image transformation:** The bounding box crops the frame to a face image. The cropped image, as the target region, is then resized, 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 anatomical landmarks (Fig. 7).

**Fig. 7:** The generated mesh topology (a), its vertices with index (b), and viewed in AR (c)**Fig. 8:** Hair segmentation module

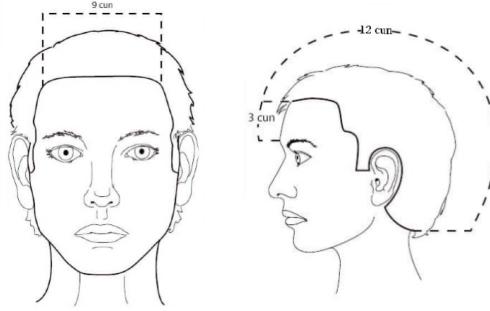
### 4.2 Hair Segmentation

According to the acupoints national standard, the center of the frontal hairline is a critical facial anatomical landmark. We adopt a pre-trained lightweight model[41] to compute the soft binary mask that separates hairs from the background. It embeds the result from the previous round as the alpha channel to the current RGB frame (Fig. 8), which accelerates the process and better supports the AR scenarios.

### 4.3 Facial Acupoints Definition

We adapt the standard B-cun method to facial acupoints localization. This method measures the length and width of each part of the body with the body surface condyles as reference landmarks and determines the position of acupoints accordingly. Then, it decides a unit Cun, a length between the set of two bone nodes divided into certain equal parts, as the basis for setting acupoints. Particularly on a face, we obtain the Cun system from the estimated hair mask and face landmarks. Then facial acupoints bank on the Cun system are shown in Figure 9. To start with, we set three facial anatomical landmarks in consonance with the national standard. In order to differentiate these three points to acupoints, we group them and name it as Channel RHD:

- **RHD1, Yintang:** The midway between the medial ends of the eyebrows.



(a) B-cun on the front face (b) B-cun on the side face

**Fig. 9:** B-cun on the head from National Standard of the People’s Republic of China, Acupoints [42]. Pictures from[43].

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

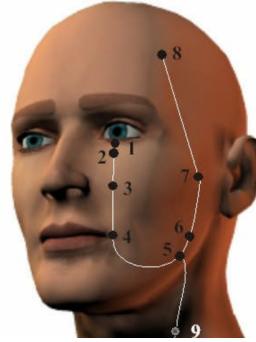
**Table 1:** Information of an example acupoint Sibai(ST2) in the data file.

- **RHD2, Middle of anterior hairline:** Intersection of anterior hairline and anterior midline.
- **RHD3, Pupils**

From Fig. 9b, the distance from RHD1 to RHD2 decides the unit Cun as  $uc = d_{RHD1 \rightarrow RHD2}/3$ . On that occasion, like RHD points definition, we could decide all facial acupoints accordingly.

For example, acupoint Sibai is by standard described as “On the face, on the infraorbital foramen, 0.5 Cun inferior to ST1 when eyes focus.” The B-cun method indicates that Sibai is vertically aligned with RHD3 (when eyes are focused) while horizontally deriving from the ST1 position. We translate this description into expressions in Table 1. The table also demonstrates what information we keep for each acupoint in the configuration file.

We group the acupoints by the meridian channels and connect those points on the same meridian channel by the flow. For example, the previously stated acupoint, Sibai, belongs to the ST (Foot’s Yang Supreme Stomach Meridian) channel. Figure 10 illustrates acupoints on the ST channel and their flow.



**Fig. 10:** Illustration of ST channel on the head. Picture from [44]

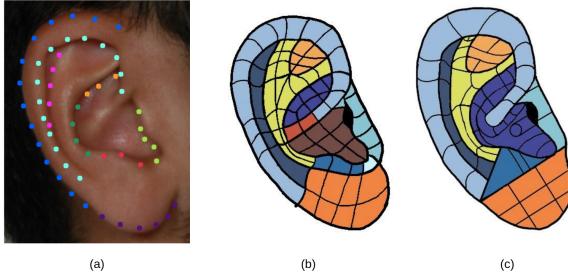
## 5 Auricular Zone Map Visualization

We started from preparing two auricular zone maps (Fig. 11 (b)(c)) according to the standards (Fig. 2 (b)(c)) as templates. Each template is described as image  $I$  with its corresponding set of landmarks  $L_I$  (Fig. 11 (a)). In order to draw auricular zone maps on frames based on anatomical landmarks, a common sequence of steps includes:

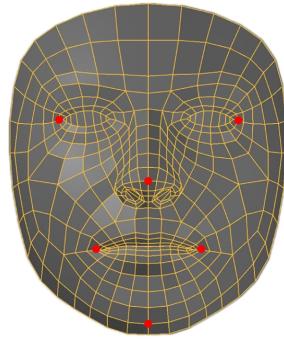
- Grab each frame as image  $J$ .
- Get the cropped ear image (Sec. 5.1) from  $J$ , each of which as  $J_c$ , and its landmarks set  $L_J$  (Sec. 5.2).
- Make  $J_c$  the same size as  $I$ , while transform  $L_J$  to  $L_I$  coordinate accordingly. Then  $L_J$  and  $L_I$  become correspondence points.
- Obtain the triangle list  $T$  from  $L_J$  (Sec.5.3).
- Warp and alpha blend image  $I$  and image  $J_c$  to get the result  $M$  (Sec.5.4).
- The target points on the morphed image  $M$  are correctly aligned.

### 5.1 Ear Detection from Facial landmarks

From the facial landmarks, we first make a quick coarse estimation of an ear bounding box. The measuring of bounding box position is to find the location of the superior attachment of the pinna. In photography, the 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 turned away slightly, such that the ear on the opposite side of the camera is just out of the shot. Even though different face and ear shapes make the angles of heads vary, the facing angle is about 45°.



**Fig. 11:** A selected image that: (1) annotated with 55 landmarks, (2) is transformed to Oleson’s auricular zone maps, and (3) is transformed to GBT13734 standard auricular zone maps.



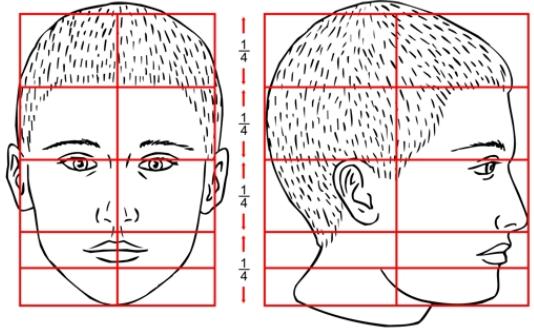
**Fig. 12:** The predicted mesh topology given by[11] annotated with the selected 6 points.

From a side view, we first estimate the head pose from a selected set of face landmarks. The pose estimation problem refers to estimating its relative orientation and position with respect to a camera and is usually described as Perspective-n-Point (PnP) problem. The goal of such a problem is to find the pose of an object with a calibrated camera,  $N$  3D object points, and the corresponding 2D projection points. We apply OpenCV `solvePnP(p3d, p2d, cM, dist)` with selected 2D and 3D landmarks as input. Kartynnik et al.[11] presents a canonical 3D facial mesh where they estimate 3D positions for vertices. Since  $N \geq 4$ , we choose 6 vertices accordingly and illustrate them on the provided canonical mesh (Fig. 12). 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  $\vec{V}(x,y,z)$  indicating the head pose towards the camera. The rules to generate ear bounding boxes are:

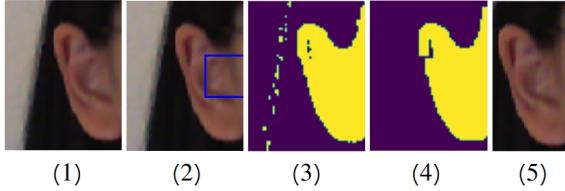
- If the rotation angle on Y-axis is within  $45^\circ$ , both ears might be seen. Otherwise, only one ear shows up.



**Fig. 13:** Proportion of the head[45]

- Rotation on X-axis determines the vertical range.  
Rotation on Z-axis controls the horizontal one.

Figure 13 illustrates the proportions of the 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 the face border to determine the coarse bounding box 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. 15) before actual landmarks localization that crops the regions of interest (ROIs) much precisely and rule out false detections. Figure 14 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 limits to the HSV frame. The lower- and upper-bound



**Fig. 14:** FineCropEar module process. (1) An input coarse cropped ear image, (2) The image patch to decide skin color boundaries, (3) The ear mask from skin color detection, (4) The mask after an opening operation, (5) The final cropped ear image.

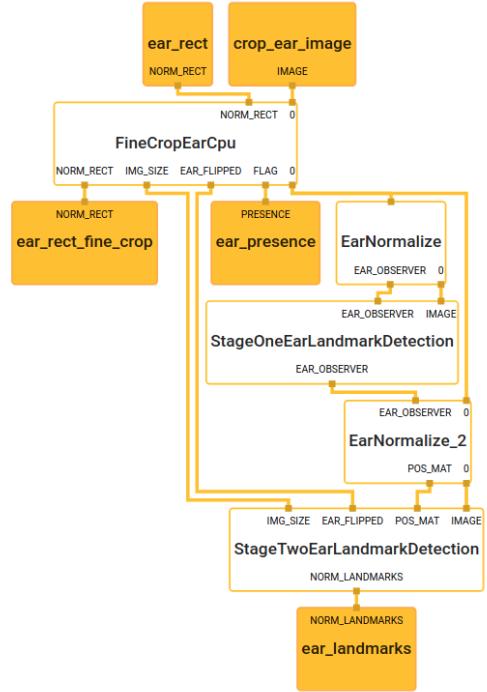
of HSV color is from the color range in a sample patch between ear and face. The patch decides the skin HSV color boundaries thus leveraging different lighting conditions and ethnicities. Then an opening operation removes small false-positive skin regions. The contour with the biggest bounding rectangle is the final target ROI.

## 5.2 Ear Landmarks localization

We utilize an unconstrained ear recognition framework from Hansley et al.[35] that works on ears-in-wild scenarios. It resolves the problem of pose variation by designing a two-stage landmark detector. We modified the original framework with supports to project the generated 55 landmarks back to the original image and construct the EarLandmarksSubgraph as shown in Figure 15. Figure 16 gives an example of our framework based on their work. Comparing landmarks results between the first stage (Fig. 16 (5)) and the second stage (Fig. 16 (7)), stage one detector is more robust to pose variations while stage two detector is more accurate for well normalized images.

## 5.3 Triangulation

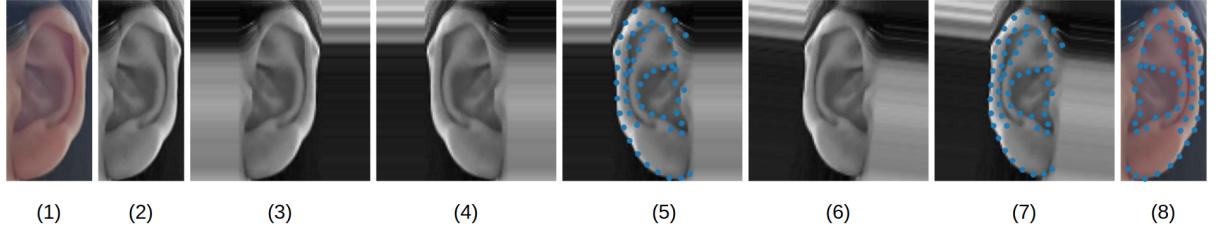
Once we have a series of anatomical ear landmarks as the point set, we need to structure the set to triangles for morphology and rendering. For a finite point set  $V = v_1, v_2, \dots, v_n$  in the Euclidean space, the point-set triangulation is a finite collection  $T$  of  $d$ -simplices of  $V$  that satisfies: (1) The union of all these simplices equals convex hull of  $V$  and (2) Any pair of these simplices intersects in a common face[46]. The most commonly used method is Delaunay Triangulation[47] that subdivides each input image into a triangular mesh. For a set of ear landmarks  $P$ , Delaunay triangulation describes the connected triangles' vertices



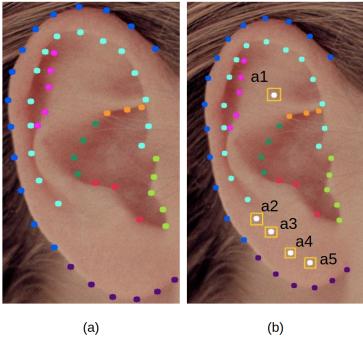
**Fig. 15:** Single ear landmarks detection Subgraph

$DT(P)$  where no point in  $P$  lies within the circumcircle of any other triangle in  $DT(P)$ . An advantage to take Delaunay triangulation is that it maximizes the smallest angles and avoids skinny triangles[48]. However, the auricular zone mapping is complex with high precision. In particular, when the landmarks are not well aligned with dividing vertices among different zones, the triangles should minimize the distortion of morphology for each auricular zone as much as possible. Based on that, we add another 5 control points derived from the original 55 landmarks. Figure 17 (a) shows a cropped ear with ground truth landmarks from “In-the-wild” ear database and its additional 5 control points (Fig. 17 (b)). Compared to Figure 4,  $a_1$  is roughly the tip of triangular fossa while  $a_2 - a_5$  are proportional dividing points between antitragus and the lobule edge.

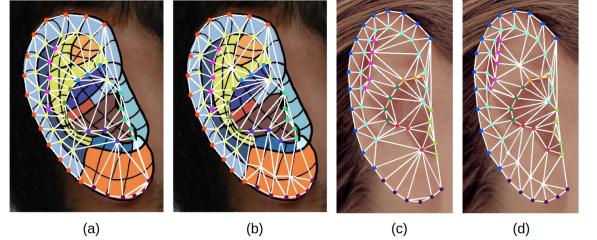
Then these triangular parts are warped and blended to produce the morph from the annotated template auricular zone maps to the detected ear image frames. Figure 18 shows the images after triangulation. Compared to triangles from Delaunay triangulation (Fig. 18 (a)), our triangles capture approximately similar regions in the recolored map (Fig. 18 (b)). Thereby,



**Fig. 16:** Example of our cropped ear landmarks localization based on Hansley et al.[35]. (1) Cropped ear image, (2) Convert the image to grayscale, (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) Back-project landmarks to (1). The original framework contains step (1)-(7) while our pipeline drops step (4) thanks to the face alignment step and adds position mapping to support (8).



**Fig. 17:** Cropped ear from “In-the-wild” ear database[40] annotated with (a) 55 ground truth landmarks, (b) additional 5 control points



**Fig. 18:** Triangulated crop ear images. Source cropped ear with landmarks and auricular zone maps (Fig. 11 (b)) that are triangulated by (a) Delaunay method and (b) our method. Sample cropped ear with ground-truth landmarks that are triangulated by (c) Delaunay method and (d) our method.

our triangle list transforms the point correspondences to region correspondences much properly for a new image (Fig. 18 (c)(d)).

#### 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  $M$  for each  $T_i$ . The amount of blending is controlled by the  $\alpha$  parameter where

$$M(x, y) = (1 - \alpha)I(x, y) + \alpha J(x, y) \quad (1)$$

where  $0 \leq \alpha \leq 1$

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

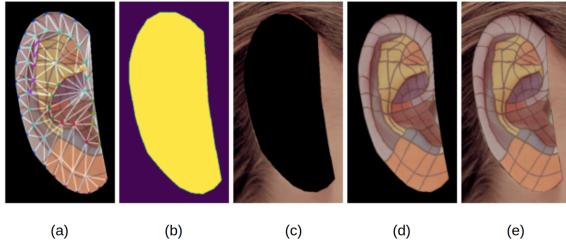
- Calculate the affine transform that maps the  $(Ti_1, Ti_2, Ti_3)$  in  $L_I$  and  $L_J$  to that in  $L_M$ .
- Transform all pixels inside the triangle of  $I$  and  $J$  to the morphed image  $M$ .

- Alpha blend the triangle patches and save to  $M$ .

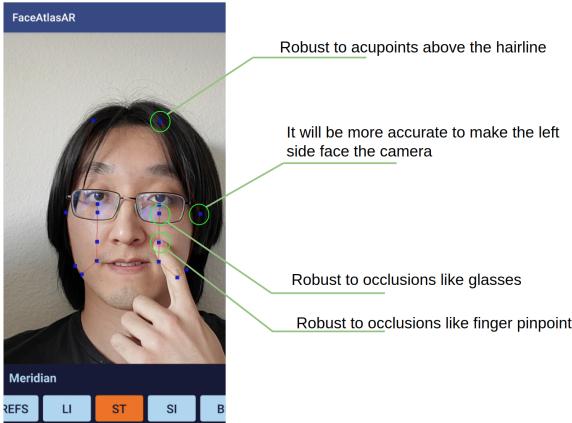
Figure 19 shows the results of applying the above technique. The morph image is a 70% blend of the input frame on the template mapping one. We see that since all the landmarks lie on the ear anatomically, the features are well aligned. Thus the triangles and the auricular zones are aligned as a result.

## 6 Experiments and Results

We designed our experiments to validate each module in the pipeline. Thus, in the following parts, we present separate results for the overall application, facial acupoints localization, and auricular zone mapping.



**Fig. 19:** Ear morphing and alpha blending. (1) Result  $M$  after morphology with re-mapped landmarks and triangles, (2) Mask of  $M$ , (3) Foreground of the morphed image, (4) Background from the input frame, and (5) Final blended result.



**Fig. 20:** A user presses an acupoint on ST meridian channel.

## 6.1 Application Screenshots

Figure 20 presents a screenshot from FaceAtlasAR Android app. We also show the visualization of acupoints grouped by meridian channels in different poses (Fig. 21). Thanks to the robustness of face alignment towards occlusion, users would not find problems pointing or pressing a target acupoint.

## 6.2 Experiments for Facial Acupoints Localization

We set up an experiment to exam the accuracy of FaceAtlasAR when localizing 4 reference points and 69 facial acupuncture points. Initially, we assign all the target positions into 3 groups by the localization complexity (see Tab. 2).

**Table 2:** Acupoints categorized by localization times

Method	Direct	One-time Proportional	Multiple Proportional
Reference points	0	4	0
Acupoints	38	16	15



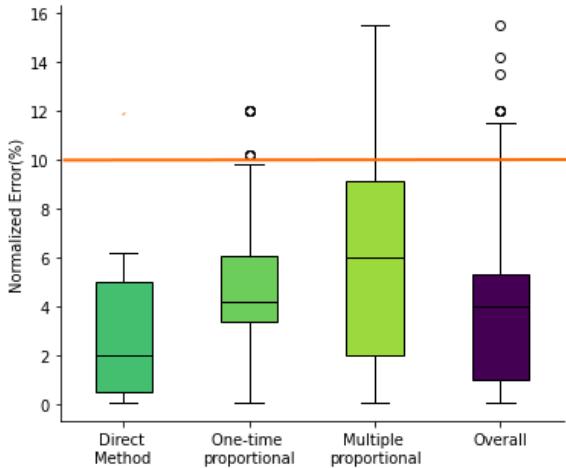
**Fig. 21:** Android application screenshots for displaying acupoints grouped by meridian system in different poses

To obtain ground-truth acupoints locations, we apply the Palpation for tenderness method (Fig. 3) in addition to the B-cun method. Therefore, instead of testing the approach on a standard face landmark dataset, we create a dataset for testing ourselves. The pictures captured in different angles make each target point get measured 5 times. Although the system tolerates different angles adequately, a target point will be more well-localized when it faces straight to the camera without any angles, which is in line with our habit of using such products. Therefore, in our experiments, we captured pictures in common usage scenarios (Fig. 21).

Then referring to the metrics of the general face landmarks datasets, we analyze the following metrics: **Normalized Mean Error(NME, %):**

$$NME(P, \hat{P}) = \frac{1}{M} \sum_{i=1}^M \frac{\|p_i - \hat{p}_i\|_2}{d} \quad (2)$$

It exams the NME between the ground truth ( $P$ ) and the predicted point ( $\hat{P}$ ) coordinates for each image respectively.  $M$  is the number of landmarks of each image;  $\hat{p}_i$  in  $\hat{P}$  is the  $i$ -th predicted landmark coordinates and  $p_i$  in  $P$  is the  $i$ -th ground truth landmark coordinates;  $d$  is the normalization factor. We



**Fig. 22:** Distribution of normalized error for all target points

adopt inter-pupil (eye to eye center) distance from the COFW dataset as the normalization factor.

#### Failure Rate (FR, %):

$$FR = \frac{1}{K} \sum_{k=1}^K [NME_k \geq 10\%] \times 100 \quad (3)$$

For one image, if NME is larger than a threshold (i.e. 10%), it is considered a failed prediction. Figure 22 shows the distributions of normalized error for all target points, including 4 reference points and 69 acupoints, with each point measured 5 times. The orange line indicates the threshold (10%) for the failed prediction. Table 3 shows the experiment results. Our approach outperforms some face landmark benchmarks because users tend to make target points facing the camera for a better localization experience, which matches the user scenarios. 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.

In summary for facial acupoints experiments:

- Our system properly displays the requested acupoints on selected meridian channels.
- The system tolerates position offsets while endures tilt and rotation to some degree.
- When the acupoints hide in one view, they will be visible and more accurate in another one. For

**Table 3:** Evaluation of target points localization

Category	NME	FR <sub>10%</sub>
Direct Method	2.64	0
One-time Proportional	4.97	9.0
MULTiple Proportional	5.07	12.5
<b>Overall</b>	<b>3.90</b>	<b>5.13</b>

example, the frontal face hides acupoints in the ear region; thus, to view them around the left ear properly, the user needs to turn the head so that the left ear faces the camera.

## 6.3 Experiments for Auricular Zone Mapping

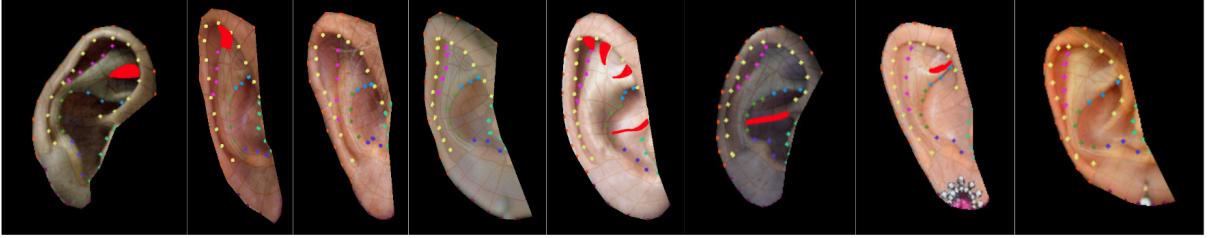
For auricular zone mapping, we evaluate our approach on two ear databases. One of them is the “In-the-wild” ear database[40] collected in the wild with ground-truth ear landmarks. The other one is AMI Ear Database[49] collected in the studio without landmarks.

**“In-the-wild” Ear Database:** Their ear images are under a challenging environment such as heavily posed angles, significant lighting variations, notable occlusions, variant resolutions, and significant aging. The whole datasets are wildly collected, while our case prefers portraits shot by front cameras. Besides, the ears in their dataset are tiny in proportion to the image, making it hard to recognize detailed structures. Therefore, we filter out the pictures if their cropped ear image is less than a threshold ( $(width + height) < 400pix$ ). With this database, we assess our system’s ability to work in different environments with tortoise-shaped ears.

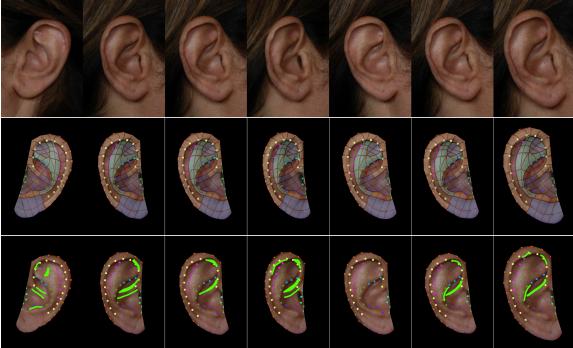
**AMI Ear Database:** The ear images are collected from 100 different subjects, each with the right ear in six poses (front, up, down, left, right, zoom) and one left ear image (back). Ground-truth ear landmarks are not available for this dataset, but their environment setup and lighting conditions are similar to our system requirements. Thus, it resembles the actual situation to explore the effects of the poses.

Despite the lack of manual gold standard, this experiment enables assessment of performance on a heterogeneous dataset. Figure 23 and Figure 24 present sample results from the two datasets. The red and green marks denote areas that failed to match the ground-truth annotations.

The graph in Figure 25 shows the number of cases on a percentage of incorrect prediction for the two datasets.



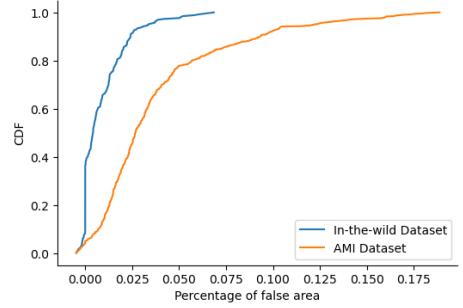
**Fig. 23:** Sample annotated results from “In-the-wild” dataset



**Fig. 24:** Sample annotated results from AMI dataset.  
Row 1 shows the original images, row 2 shows the resulting auricular zone maps, and row 3 labels failure areas of the results with green marks.

It appears that performance on “In-the-Wild” datasets is much better than that on the AMI one. Out of 258 cropped ears in the “In-the-Wild” dataset, 67% images are accepted as completely correct mapping. Most incorrect mapping only affects up to 3% of the whole ear, meaning that minor incorrect mapping in a zone may not affect other zones. For results on the AMI dataset, the percentage of false mapping is 2% – 12%, which indicates that offset labeling in a zone will ruin the adjacencies. Thus both areas are annotated as incorrect. Example such as Figure 24 (1)(2)(4). Compared to the results from the two datasets, the ground-truth anatomical landmarks provide accurate morphology.

Initial comparison between the two datasets shows that the system is robust to the lighting environment and occlusion to some degree. Comparably, poses are more inclined to decline the performance. Based on the observations, we further evaluate the accuracies on the AMI dataset categorized by 6 poses. Figure 26 box plots the percentage of false area by categories. The

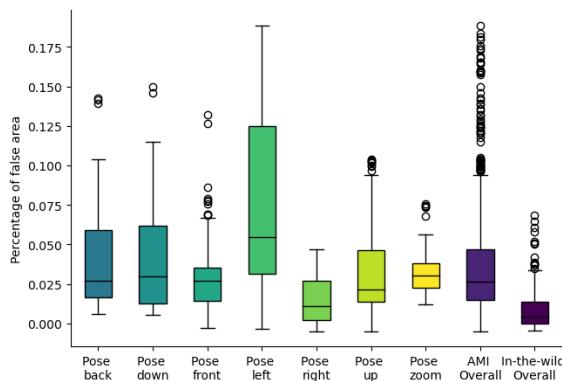


**Fig. 25:** Cumulative distribution for (a) “In-the-Wild” dataset and (b) AMI dataset.

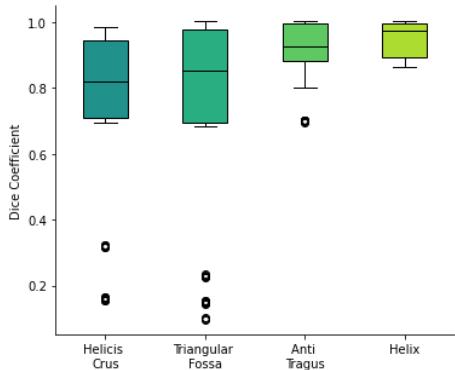
back pose is a left-side profile while others are right-side profiles with the individual facing forward (front), looking up and down (up, down) and looking left and right (left, right). The zoom category is also the right profile with the subject facing forward but with a different camera focal length. Ideally, when users closely exam the right ear, they face the right ear straight to the front camera as much as possible. Thus, the right pose resembles the real settings the most. Other poses simulate occasions that are also likely to happen. The left pose relates to when both ears appear in the camera, making each of them partially occluded.

From the results, the idea pose (right) suppresses most of the false percentages below 0.03 (3%), whereas the LEFT pose performs the worst (0.05 – 0.18) because the ears are usually not facing the cameras enough to get the model well-aligned. The system tolerates rotations on roll and pitch axes, controlling the false area within 0.08.

Results on AMI indicate that without ground-truth landmarks, structures such as the Concha Cymba, Helicis Crus, and the triangular fossa are more likely to fail to match due to a lack of specific anatomical landmarks. Hence, we further adopt dice coefficients to measure some structures of interest from the right



**Fig. 26:** Percentage of false area by categories. From left to right: 7 pose categories from AMI dataset, overall results in AMI dataset, and overall results in “In-the-Wild” dataset



**Fig. 27:** Dice coefficients for Helicis Crus, Triangular Fossa, Anti Tragus, and Helix.

pose images (Fig. 27):

$$\text{Dice Coefficient} = \frac{2TP}{2TP + FN + FP} \quad (4)$$

where TP is the true positive, FP is the false positive, and FN is the false negative area. We use this measurement to focus on deep structures such as the Helicis Crus, Triangular Fossa, Anti Tragus, and Helix, which failed through our initial experiment due to lack of specific anatomical landmarks. Results show that more complex structures have higher rates of false estimation.

## 7 Conclusion

In this paper, we proposed FaceAtlasAR, an end-to-end facial and ear acupoint localization AR system. It overlays auricular zone maps and facial acupoints to users interactively in real-time. The high accuracy of the estimation and the robustness empower users with little experience in acupuncture to pinpoint acupoints. Future research should consider the potential effects of the proportional method more carefully. For example, multiple times calculation declines the localization accuracies. Future research is needed to investigate 3D interactions in different user scenarios. For instance, we could track users’ hands/fingers while interacting with a target acupoint. Otherwise, external hardware like acupuncture needles may be applied for interaction in a more precise way.

## Declarations

- Funding: No funding was received to assist with the preparation of this manuscript.
- Conflicts of interest/Competing interests: The authors declare they have no financial or non-financial interests.
- Availability of data and material: The data that support the findings of this study including public datasets AMI dataset and “In-the-wild” dataset, which are available in the public domain. The data citation in the main article has the full URL. Other data are available on request from the corresponding author, M.Z.
- Code availability: Code will be available at <https://github.com/ZhangMenghe/mediapipe/>
- Ethics approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.
- Consent to participate: Informed consent was obtained from all individual participants included in the study.
- Consent for publication: The authors affirm that human research participants provided informed consent for publication of the images in Figure(s) 7, 8, 20 and 21.

## References

- [1] Kaptchuk, T.J.: Acupuncture: theory, efficacy,

- and practice. *Annals of internal medicine* **136**(5), 374–383 (2002)
- [2] Li, T., Tang, B.Q., Zhang, W.-B., Zhao, M., Hu, Q., Ahn, A.: In vivo visualization of the pericardium meridian with fluorescent dyes. *Evidence-Based Complementary and Alternative Medicine* **2021** (2021)
- [3] Berman, B.M., Lao, L., Langenberg, P., Lee, W.L., Gilpin, A.M., Hochberg, M.C.: Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: a randomized, controlled trial. *Annals of internal medicine* **141**(12), 901–910 (2004)
- [4] Yuan, J., Purepong, N., Kerr, D.P., Park, J., Bradbury, I., McDonough, S.: Effectiveness of acupuncture for low back pain: a systematic review. *Spine* **33**(23), 887–900 (2008)
- [5] Vickers, A.J., Vertosick, E.A., Lewith, G., MacPherson, H., Foster, N.E., Sherman, K.J., Irnich, D., Witt, C.M., Linde, K., Collaboration, A.T., et al.: Acupuncture for chronic pain: update of an individual patient data meta-analysis. *The Journal of Pain* **19**(5), 455–474 (2018)
- [6] Azuma, R.T.: A survey of augmented reality. *Presence: teleoperators & virtual environments* **6**(4), 355–385 (1997)
- [7] Li, F., He, T., Xu, Q., Lin, L.-T., Li, H., Liu, Y., Shi, G.-X., Liu, C.-Z.: What is the acupoint? a preliminary review of acupoints. *Pain Medicine* **16**(10), 1905–1915 (2015)
- [8] Huber, P., Hu, G., Tena, R., Mortazavian, P., Koppen, P., Christmas, W.J., Ratsch, M., Kittler, J.: A multiresolution 3d morphable face model and fitting framework. In: *Proceedings of the 11th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* (2016)
- [9] Chen, Y.-Z., Maigre, C., Hu, M.-C., Lan, K.-c.: Localization of acupoints using augmented reality. In: *Proceedings of the 8th ACM on Multimedia Systems Conference*, pp. 239–241 (2017)
- [10] Lan, K.-C., Li, G.-S., Zhang, J.-X.: Toward automated acupressure therapy. In: *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*, pp. 384–385 (2018)
- [11] Kartynnik, Y., Ablavatski, A., Grishchenko, I., Grundmann, M.: Real-time facial surface geometry from monocular video on mobile gpus. *arXiv preprint arXiv:1907.06724* (2019)
- [12] Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M.G., Lee, J., et al.: Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172* (2019)
- [13] AACP: Evidence Based Acupuncture Training, Acupuncture in Physiotherapy. The Acupuncture Association of Chartered Physiotherapists (AACP)s. The Acupuncture Association of Chartered Physiotherapists (AACP)s
- [14] Nogier, P.: Le pavillon de l'oreille. zones et points réflexes. *Bulletin de la Société d'Acupuncture* (20), 51–57 (1956)
- [15] Gori, L., Firenzuoli, F.: Ear acupuncture in european traditional medicine. *Evidence-Based Complementary and Alternative Medicine* **4**(S1), 13–16 (2007)
- [16] Oleson, T.D., Kroening, R.J., Bresler, D.E.: An experimental evaluation of auricular diagnosis: the somatotopic mapping of musculoskeletal pain at ear acupuncture points. *Pain* **8**(2), 217–229 (1980)
- [17] Oleson, T., Flocco, W.: Randomized controlled study of premenstrual symptoms treated with ear, hand, and foot reflexology. *Obstetrics and Gynecology* **82**, 906–911 (1993)
- [18] Weintraub, M.I., Mamtani, R., Micozzi, M.S.: Complementary and Integrative Medicine in Pain Management. Springer, ??? (2008)
- [19] Yao, J., Chen, L., Zhang, L., Zhou, S., Zheng, Q., Feng, X., You, X., Zhang, L., Li, Y.: Effect of auriculotherapy and intervention types on weight control: a systematic review and meta-analysis protocol. *Medicine* **98**(34) (2019)

- [20] Wu, C., Liu, P., Fu, H., Chen, W., Cui, S., Lu, L., Tang, C.: Transcutaneous auricular vagus nerve stimulation in treating major depressive disorder: a systematic review and meta-analysis. *Medicine* **97**(52) (2018)
- [21] Standard, C.: Nomenclature and location of auricular points. Standard, Chinese Acupuncture Society, China (April 2008)
- [22] Nogier, P.M., Cousino, A., Graff, M.: From Auriculotherapy to Auriculomedicine. Maisonneuve, ??? (1983)
- [23] White, A., Ernst, E.: A brief history of acupuncture. *Rheumatology* **43**(5), 662–663 (2004)
- [24] Kanehira, R., Shoda, A., Yagihashi, M., Narita, H., Fujimoto, H.: Development of an acupuncture training system using virtual reality technology. In: 2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery, vol. 4, pp. 665–668 (2008). IEEE
- [25] Yau, E.: Virtual Reality Helps Chinese Medicine Students Learn Acupuncture and Doctors Treat Cancer. <https://www.scmp.com/lifestyle/health-wellness/article/2169092/virtual-reality-helps-chinese-medicine-students-learn>
- [26] Godson, D.R., Wardle, J.L.: Accuracy and precision in acupuncture point location: a critical systematic review. *Journal of acupuncture and meridian studies* **12**(2), 52–66 (2019)
- [27] Molsberger, A., Manickavasagan, J., Abholz, H., Maixner, W., Endres, H.: Acupuncture points are large fields: the fuzziness of acupuncture point localization by doctors in practice. *European Journal of Pain* **16**(9), 1264–1270 (2012)
- [28] Hong, Y., Shang, H., Yang, H., Kong, Q., Wang, M., Zhang, Q.: A 3d recognition and projection system for meridians and acupoints. In: 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1357–1363 (2017). IEEE
- [29] Bulat, A., Tzimiropoulos, G.: How far are we from solving the 2d & 3d face alignment problem?(and a dataset of 230,000 3d facial landmarks). In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1021–1030 (2017)
- [30] Sagonas, C., Tzimiropoulos, G., Zafeiriou, S., Pantic, M.: A semi-automatic methodology for facial landmark annotation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 896–903 (2013)
- [31] Belhumeur, P.N., Jacobs, D.W., Kriegman, D.J., Kumar, N.: Localizing parts of faces using a consensus of exemplars. *IEEE transactions on pattern analysis and machine intelligence* **35**(12), 2930–2940 (2013)
- [32] Le, V., Brandt, J., Lin, Z., Bourdev, L., Huang, T.S.: Interactive facial feature localization. In: European Conference on Computer Vision, pp. 679–692 (2012). Springer
- [33] Zhu, X., Ramanan, D.: Face detection, pose estimation, and landmark localization in the wild. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp. 2879–2886 (2012). IEEE
- [34] Emeršič, Ž., Štruc, V., Peer, P.: Ear recognition: More than a survey. *Neurocomputing* **255**, 26–39 (2017)
- [35] Hansley, E.E., Segundo, M.P., Sarkar, S.: Employing fusion of learned and handcrafted features for unconstrained ear recognition. *IET Biometrics* **7**(3), 215–223 (2018)
- [36] Emeršič, Ž., Meden, B., Peer, P., Štruc, V.: Evaluation and analysis of ear recognition models: performance, complexity and resource requirements. *Neural computing and applications* **32**(20), 15785–15800 (2020)
- [37] Frejlichowski, D., Tyszkiewicz, N.: The west pomeranian university of technology ear database—a tool for testing biometric algorithms. In: International Conference Image Analysis and Recognition, pp. 227–234 (2010). Springer
- [38] Raposo, R., Hoyle, E., Peixinho, A., Proen  a, H.:

- Ubear: A dataset of ear images captured on-the-move in uncontrolled conditions. In: 2011 IEEE Workshop on Computational Intelligence in Biometrics and Identity Management (CIBIM), pp. 84–90 (2011). IEEE
- [39] Kumar, A., Wu, C.: Automated human identification using ear imaging. *Pattern Recognition* **45**(3), 956–968 (2012)
- [40] Zhou, Y., Zaferiou, S.: Deformable models of ears in-the-wild for alignment and recognition. In: 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), pp. 626–633 (2017). IEEE
- [41] Tkachenka, A., Karpiak, G., Vakunov, A., Kartynnik, Y., Ablavatski, A., Bazarevsky, V., Pitsarchyk, S.: Real-time hair segmentation and recoloring on mobile gpus. arXiv preprint arXiv:1907.06740 (2019)
- [42] Unknown: National Standard of the People's Republic of China, Acupoints. <https://musculoskeletalkey.com/ynsa-basic-points/> (2005)
- [43] Unknown: YNSA Basic Points. <https://musculoskeletalkey.com/ynsa-basic-points/> (2017)
- [44] Unknown: Atlas Of Acupuncture Points. [www.AcupunctureProducts.com](http://www.AcupunctureProducts.com) (2007)
- [45] Artyfactory: THE PROPORTIONS OF THE HEAD. <https://www.artfactory.com/portraits/pencil-portraits/proportions-of-a-head.html>
- [46] Rambau, J., Leal, F.S.: Triangulations of point sets (2003)
- [47] Lee, D.-T., Schachter, B.J.: Two algorithms for constructing a delaunay triangulation. *International Journal of Computer & Information Sciences* **9**(3), 219–242 (1980)
- [48] Mount, D.M.: Cmsc 754 computational geometry. Lecture Notes, University of Maryland, 1–122 (2002)
- [49] Esther Gonzalez, L.A., Mazorra, L.: AMI Ear Database. [https://ctim.ulpgc.es/research\\_works/](https://ctim.ulpgc.es/research_works/)
- ami\_ear\_database/ (2008)