

# 基于互动镜面的智能问候系统

---

**摘要：**镜子在人们的生活中可谓无处不在，人们用其整理仪表、装饰房间，稍稍设计还可美化生活。近年来“互动镜面”的诞生为传统的广告、科技展览等领域注入了新鲜的活力。所谓互动镜面，是一种既可以照镜子，又可以呈现高画质投影图像的“魔镜”，与此同时用户还可通过触摸来与其进行互动。基于现已较为成熟的镜面互动技术，配以语音控制识别技术、人脸定位及跟踪技术、图像分割及识别等技术，有希望研制出一种新型智能问候系统——即在满足用户日常使用的基础之上，可以简单判断用户心情、健康状况并给出针对性建议与提醒、对用户提出的问题（如询问天气、周边路况）进行回答等。此外通过指纹识别技术，此系统将确保用户的个人隐私安全，真正成为一个全面、安全的智能系统，一面智能且贴心的“魔镜”。

**关键字：**互动镜面，语音控制识别，图像识别，智能，问候，安全

**Abstract:**

Mirrors are indispensable in our life, with which people can beautify themselves as well as their room space. And in recent years, the birth of “interactive mirror” has injected fresh vitality to the traditional advertising and science-tech exhibition. This so-called “interactive mirror” is a magic mirror, which, on top of the usage of a common mirror, can present HD projections and meanwhile interact with users through touch.

This paper is based on the relatively mature technology defined as “Specula Interactive Technology”. In the author’s opinion, a new intelligent greeting system can be developed with speech recognition technology, face detection & tracking technology, and image segmentation & recognition technology etc. This interactive mirror greeting system, apart from the basic daily use, can briefly judge a user’s mood and health status, and give corresponding advice and reminders. In addition, this intelligent system is able to answer users’ questions on weather, traffic,

reservations and so on, like Siri does on iphone. With the fingerprint recognition technology, this system will truly become a comprehensive and intelligent system with users' personal privacy totally secured; A “mirror” empowered by so many high technologies is sure to be a magic mirror, and it is very hopeful that such greeting system can be developed.

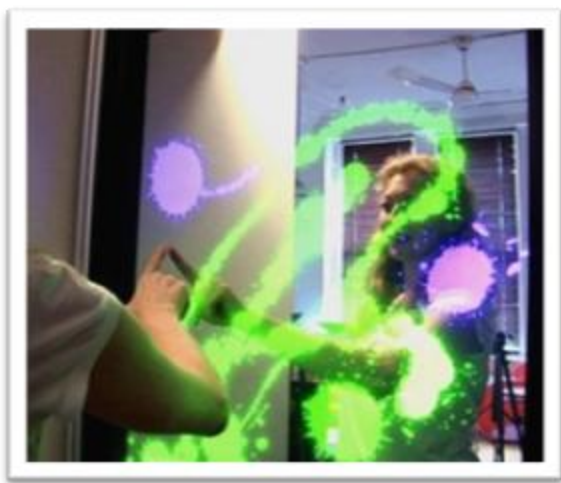
**Keywords:** Interactive mirror, speech Recognition, image segmentation & recognition, Intelligent, safe, considerable

## 目录

第一章引言 .....	5
1.1 创意来源.....	5
1.2 具体阐述.....	6
第二章 组成部分 .....	7
2.1 互动镜面 .....	7
2.2 智能问候系统 .....	8
第三章 相关技术 .....	9
3.1 指纹识别技术 .....	10
3.2 人脸定位及识别技术.....	11
3.2.1 人脸定位 <sup>[2]</sup> .....	11
3.2.2 人脸跟踪 <sup>[2]</sup> .....	11
3.3 图像分割与识别技术.....	13
3.3.1 图像分割.....	14
3.3.2 图像识别.....	16
3.4 语音控制识别技术 .....	17
3.5 温湿度感应技术.....	19
第四章 结束语 .....	20

## 第一章引言

### 1.1 创意来源



在电子技术迅速发展的今天，人们的生活越来越多地和虚拟、智能结合起来。对于传统的居家及生活用品，我们不仅需要其美观且实用，并希望它们能够变得“智能”、懂得“察言观色”，即善解人意（如智能手机的语音控制技术）。本文的创意来源于生活中常见的镜子。

镜子的用处之一便是帮助人们整理仪容。然而当你早晨起床洗漱时，如果你的镜子告诉你“你似乎没睡好，要注意身体”时，是不是觉得很贴心呢？当你要参加会议、或是前去约会时，它会告诉你什么样的着装更为得体；当你想知道天气却又懒得开电脑翻手机时，只需张口一句话，你的镜子就会告诉你近期的天气预报、甚至周边的路况；当你一脸沮丧地站在它面前时，它会善解人意地播放舒缓的音乐，甚至讲个笑话……本文所指的智能问候系统可以让这一切成为现实。

## 1.2 具体阐述

本文的创意主要分为两部分，一是“互动镜面”，二是“智能问候系统”。受到全息投影以及 Siri 等语音控制识别技术的启发，作者便想将二者结合起来，研制出一个以互动镜面为载体的智能问候系统。

通过查找相关资料，作者了解到近年来镜面互动技术日趋成熟，早已大量运用于科技展览、广告宣传、舞台效果以及新产品发布等领域。然而此项技术似乎对人们的居家生活涉及不多。于是作者思考，如果能将其引进人们家中，替代传统平面镜，是否既可满足人们的日常使用需求，也可通过人机交互技术使人们的生活变得更智能呢？

然而市面上已有的互动镜面并不能满足作者的预期：这面镜子应该更具人性化，甚是可以达到“善解人意”的程度。这让人联想到苹果公司 2011 年与 iPhone 4S 一同推出的 Siri 语音控制识别应用。Siri 不仅能针对用户语音指令做出相应的操作，还能够不断学习新的声音和语调，提供对话式的应答——而这正是作者所想要的。



与此同时，人脸识别技术，图像识别、分类及预处理技术的日渐成熟。通过图像采集和处理可以将用户的脸部特征（如表情、眼部肌肉颜色等）与相应的健康状况匹配及精神面貌相匹配，达到简单判断用户身体、心理健康状况的地步。

作者认为，若能将上述两方面相结合，此系统便可在简单判断用户健康状况、精神面貌的基础上给出相应的建议及“问候”，让用户在日常洗漱的同时可以感受到贴心的温暖，这便是作者所谓的“智能问候系统”。同时还将与指纹识别技术相结合，以保证用户隐私安全。

## 第二章 组成部分

### 2.1 互动镜面



互动镜面又称“互动魔镜”，用特殊材质镜面玻璃作为投影介质。不仅可以照镜子，同时还能播放高画质投影的图像。用户通过手指触摸镜子表面，可以实现人机交互，可谓“超大平板电脑”。而互动镜面系统是以背投方式的一种超媒体互动形式。镜面是一块高性能的双面显示屏幕，它具有视觉平方所有产品在清晰显像、耐光亮等优点。同时，它又是一面镜子。

通过综合衍射图（HOLOGRAM T.）技术与魔术光学（MAGIC OPTICS）镜面涂层完美结合，成就了拥有高贵魅力的魔幻镜面系列，也就是本文所提到的“互动镜面”，而本系统正是以互动镜面为基础的。

## 2.2 智能问候系统

本系统的主要硬件包括了高清摄像头、音频采集器、指纹识别器（指纹采集头），计算机，音箱，温湿度传感器；软件主要由指纹采集配套软件、人脸定位跟踪算法、图像分割及识别算法等等构成。

系统工作步骤为：首先用户需进行指纹识别以启动系统，确保自身隐私安全，随后高清摄像头及温湿度感应器启动，进行用户人脸定位及跟踪；用户面部定位成功后摄像头将跟踪拍摄一组高清的用户面部照片；接着选取清晰度最高的照片经过一系列处理得到相应数字信息，并交由计算机判断关于用户健康、精神面貌的各项数据；判断完成系统将对用户进行语音反馈，包括“提醒”、“信息查找”、“智能问候”、“给出建议”等等；如用户继续发出语音指令，则计算机继续采集、编译命令。

当用户离开时，即摄像头无法持续捕捉用户面部时，系统将自动进入休眠状态，摄像头停止工作。如需再次启动，用户需重新进行指纹识别。

图4 为本系统的工作流程图：

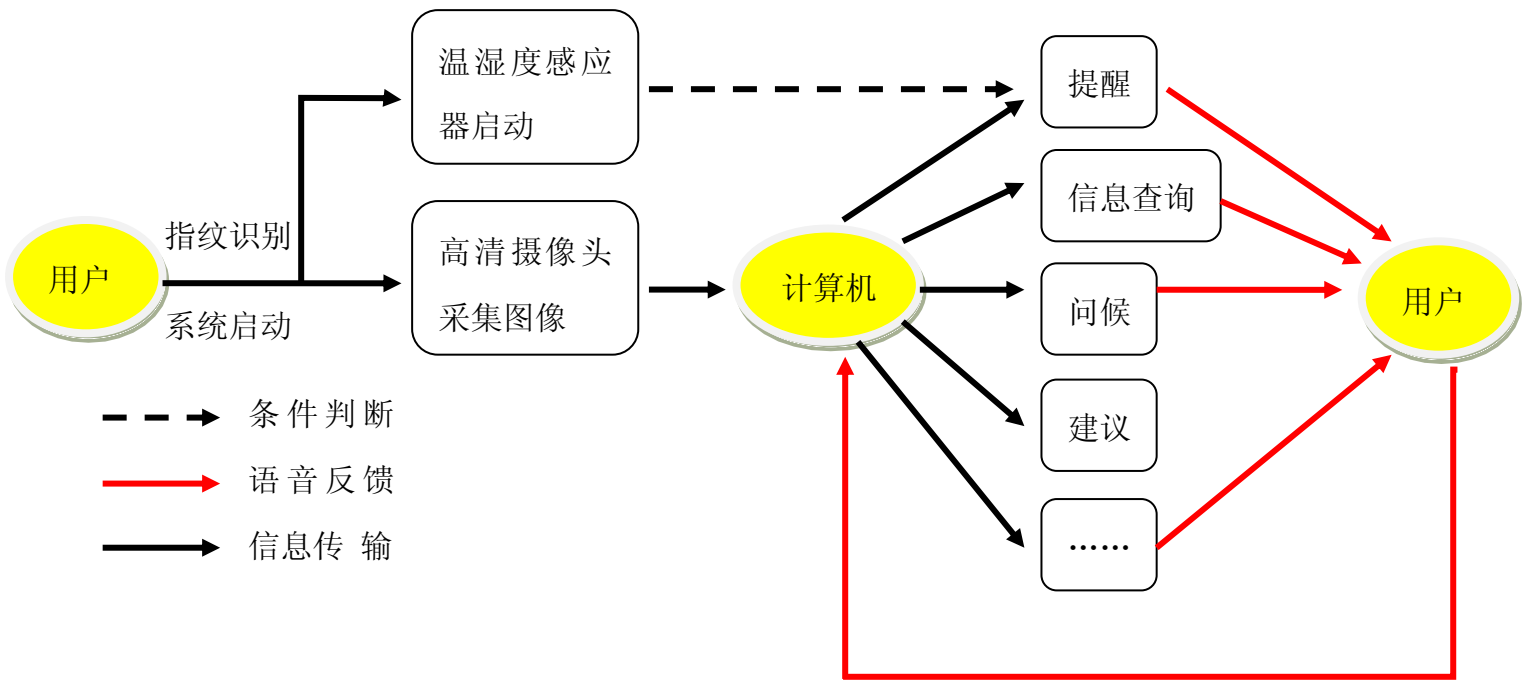
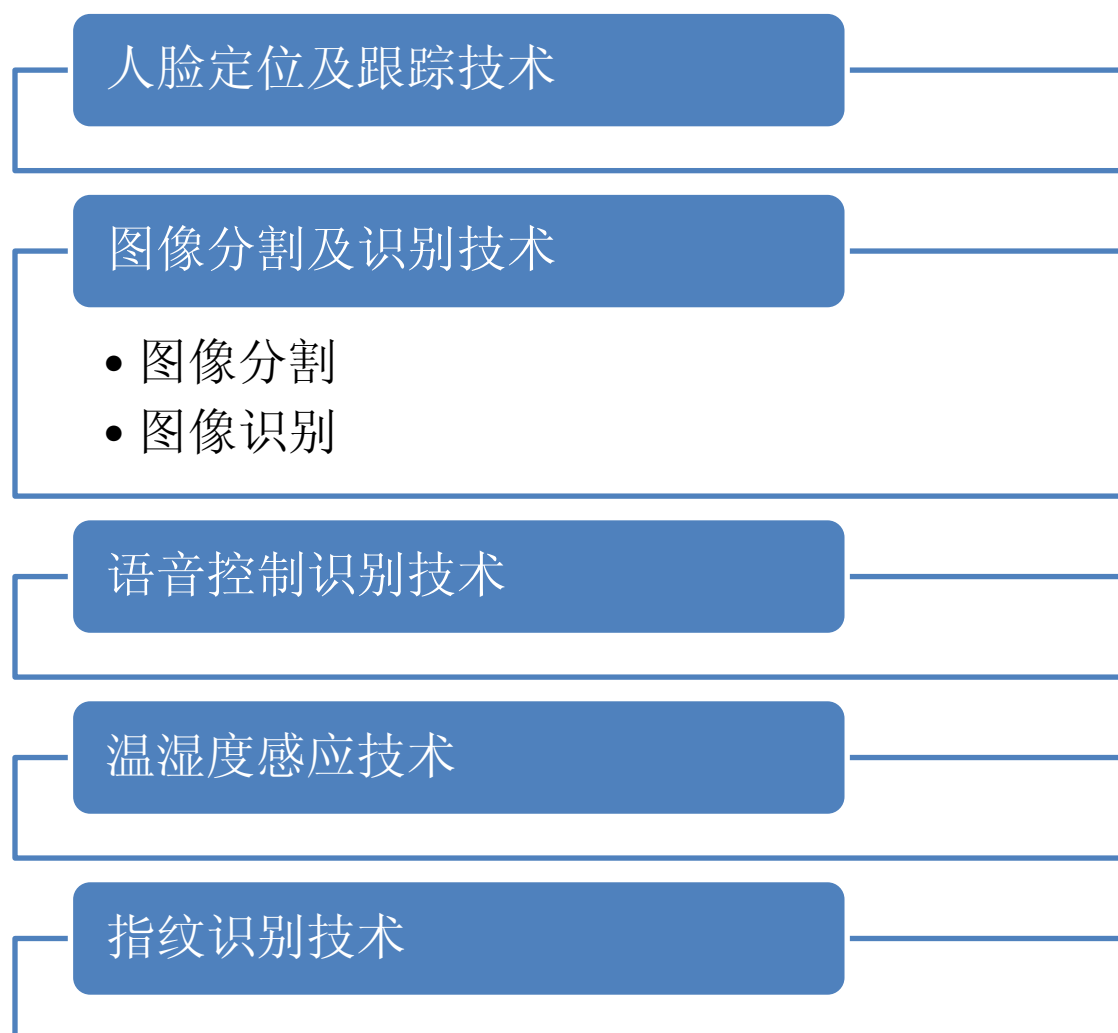


图4



### 第三章 相关技术

本系统所涉及的主要技术如下



### 3.1 指纹识别技术

正如本文摘要所述，鉴于本系统的适用场所多为卫生间、浴室以及卧室，则用户的隐私安全不言而喻。经过查找资料，由于语音身份识别技术处理速度慢、分辨率差、目前尚未成熟等等，本系统采用的是发展相对完善、成熟的指纹识别技术。

由于本系统主要用于家用，用户数量较少，所以与检测自动指纹识别系统（AFIS）相比，本系统的指纹识别可以省去庞大的数据库，从而提高运算处理速度。用户首次使用本系统时，需要将指纹录入系统并上传服务器，使用过程中，用户数据可由用户自行更改。大致流程<sup>[1]</sup>如下：

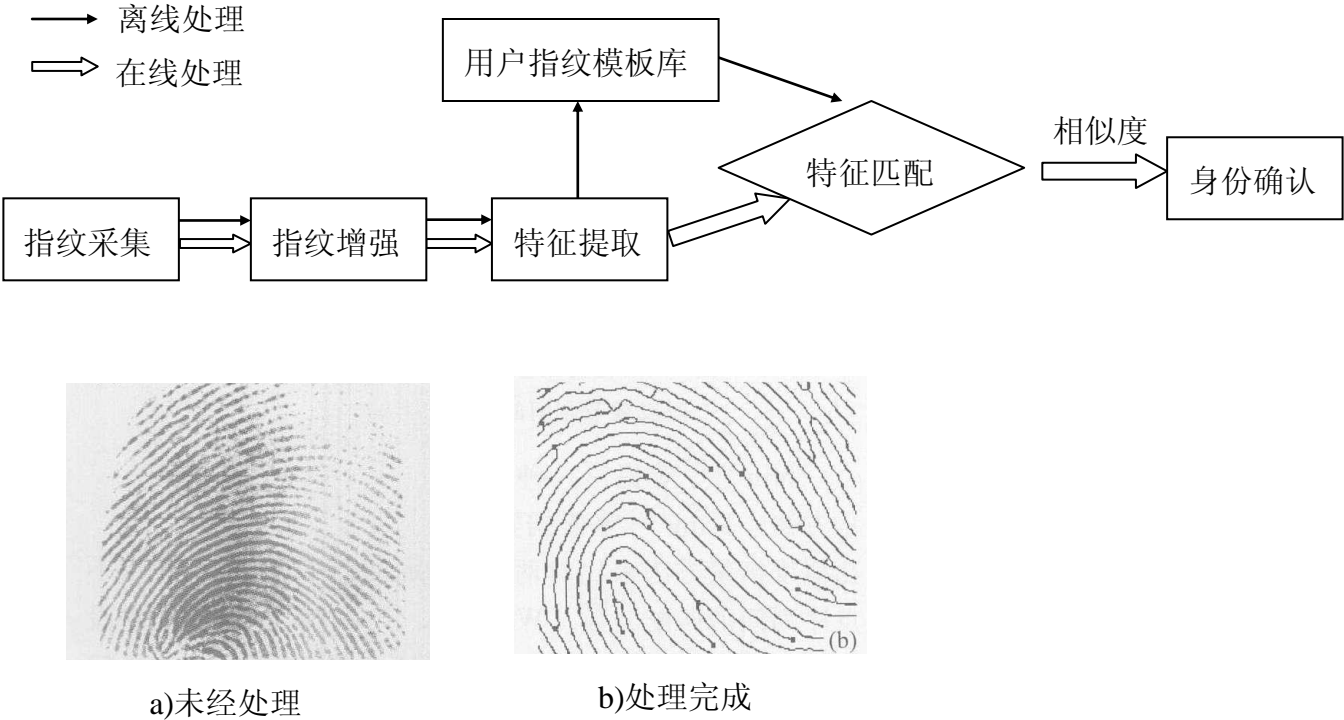
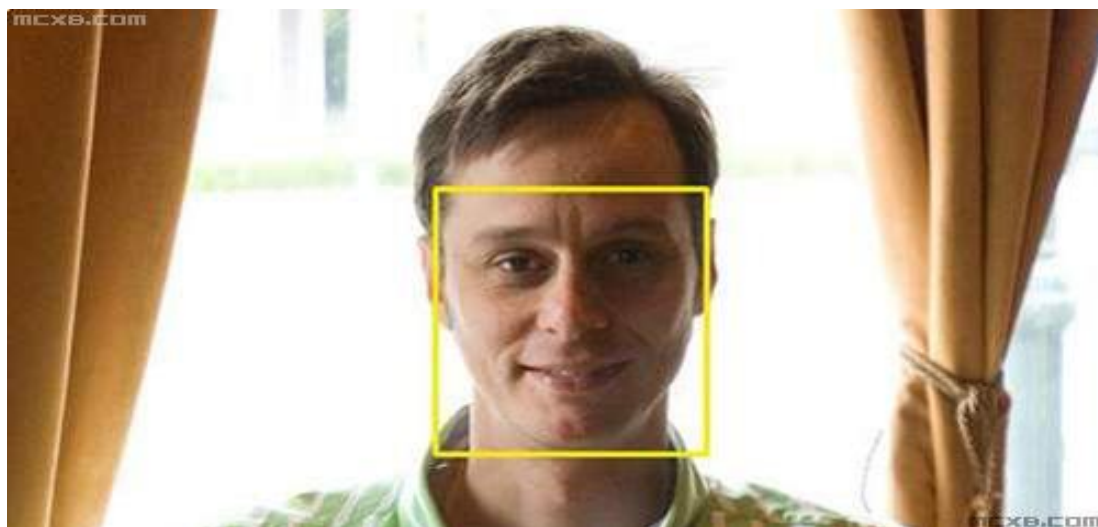


图 6

近十年内，在线指纹采集技术和自动指纹识别算法取得了显著的发展，通过指纹分割、归一化、方向场计算、脊线滤波及映射、细节点提取和指纹匹配等技术（在作者所引用文献中以上技术均已证实，在此无需赘述），用户的身份难以造假，这在很大程度上确保了本系统的安全性与可靠性。

### 3.2 人脸定位及识别技术



本节的主要目的是通过镜子上方的摄像头对用户的面部进行定位与跟踪，在确定用户面部具体位置之后，摄像头将对其进行拍摄一组高清晰照片，以进行接下来的处理。

本节涉及到的主要技术分为两方面，一是人脸定位，而是人脸跟踪。接下来具体方法简述如下：

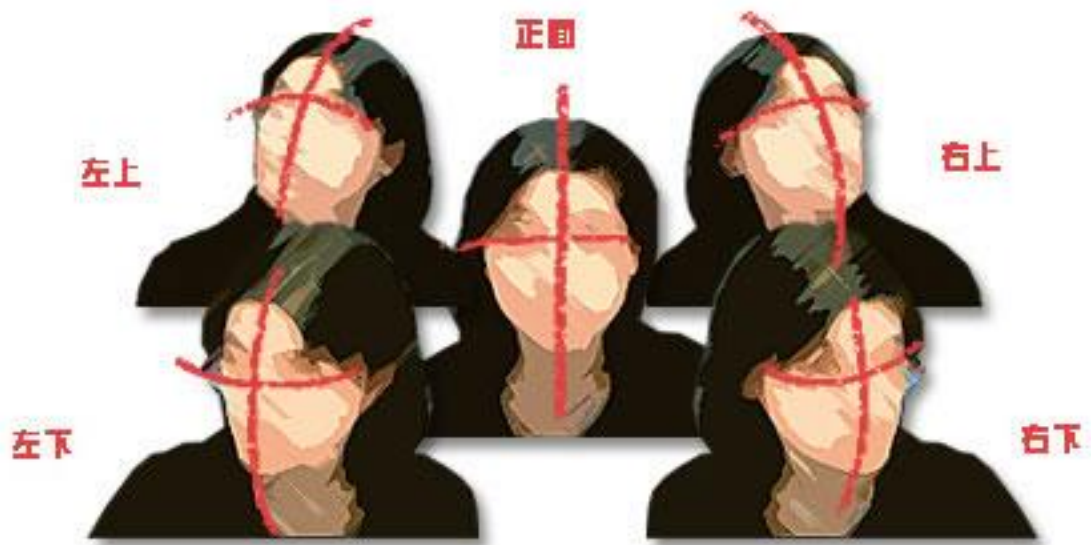
#### 3.2.1 人脸定位<sup>[2]</sup>

采用帧差法将连续的两帧视频图像进行预处理，去除掉背景中不动的部分，将处理的范围缩小到运动的区域，然后结合人脸肤色模型对人脸进行粗检测，经过这一步就可以将人脸大致定位出来。

为了增加定位的精度以及后续跟踪的准确率，在经过该步处理后的肤色区域中，利用基于统计学的 adaboost 算法对人脸进行精确定位，将人脸用外接矩形框标记出来。

#### 3.2.2 人脸跟踪<sup>[2]</sup>

在第二部分也就是人脸的跟踪部分，有效地结合了第一部分人脸定位中所采用的算法结果，采用“将 adaboost 人脸检测算法和 Camshift 人脸跟踪算法相结合”的方案对人脸进行跟踪。将在单帧图像中经过 adaboost 算法处理所得到的人脸外接矩形框的大小、位置作为 Camshift 搜索窗口的初始值，在跟踪下一帧图像时，只需要在该搜索窗口稍大的区域进行人脸的搜索即可。



本文在人脸检测部分选择引用帧差法、人脸肤色模型法与人脸检测 adaboost 算法相结合的算法, 以此创新性地添加检测旋转任意角度人脸的辐射模板法。已有相关实验验证, 可以在保证人脸检测精确率的前提下大幅度地提高人脸检测的效率, 并且对旋转人脸有很好的检测效果。具体实现方法于参考文献中已有详细说明, 本文在此同样不予赘述。

### 3.3 图像分割与识别技术

当镜面上方的摄像头对用户面部进行定位、跟踪之后，便会拍摄一组高清晰的面部照片，并筛选出清晰度最高的一张进行接下来的一系列处理工作——这主要涉及到图像分割和图像识别两种技术。

经过处理，用户面部照片会被分为眼角、下眼睑、嘴角、嘴唇、两颊等几个部分，并且分别得到其对应的数字信号。本系统的主要功能之一便是“对用户的健康状况做出简单判断”，故系统内部需预设相应参数使之与用户真实情况进行比较。如：



黑眼圈——疲劳



嘴角向下——沮丧

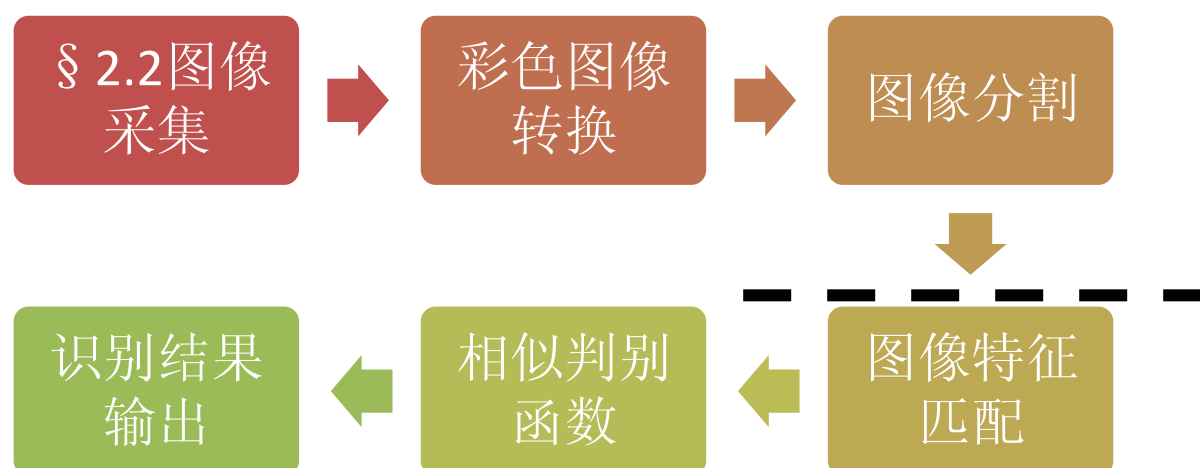


嘴唇干裂——缺水

.....

通过一系列流程将采集得到的数据与系统预先设定的数据进行比较，便可将用户的健康状况与面部特征对应起来，继而通过前文所提到的“语音控制识别技术”给出建议。

下图 是本环节的具体流程



### 3.3.1 图像分割

#### • 彩色图像转换

本文提出的图像分割主要可分为两步：一、将彩色图像转换为灰度图像；二、通过直方图阈值化，分别得到每部分的最佳阈值后对灰度图像进行分割。

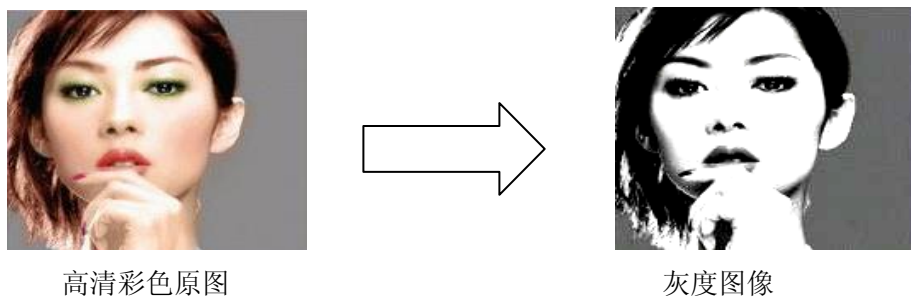
由彩色图像转换为灰度图像有多种方法，如 MATLAB 编程、HLS 模型转换法等。

通过查找文献，作者认为选用 HLS 模型转化法<sup>[3]</sup>所得到的灰度图像质量更高，现简述如下：

H、L、S 分别代表色调、亮度和饱和度。该模型与一般的色调、亮度、饱和度模型不同。一般的色调、亮度、饱和度是从色度学角度与 RGB 模型、CMYK 模型等模型之间的线性转换而来，它们的转换是通一线性转换矩阵来实现的。而 HLS 模型是基于视觉感受的模型，它的转换算法是非线性的，其转换的算法为：

$$Y = \frac{[\text{MAX}(R, G, B) + \text{MIN}(R, G, B)]}{2} \times 3 \quad (1)$$

(1) 式中：Y 为像素点的亮度，R，G，B 分别为三基色的相对强度。采用该算法得到的效果图如下：



#### • 最大熵阈值法分割

通常的图像分割<sup>[4]</sup>方法有最大熵阈值法、边缘检测法、神经网络法等等。得到灰度图像后，本文采用最大熵阈值法对用户面部照片进行分割。

最大熵阈值化方法是一种自动的非参数、无监督的阈值选择法，它是基于“分割后物体与背景灰度分布的信息熵之和为最大”的优化判别式、测度准则来选择图像分割的阈值，即最佳的阈值在该测度函数取得最大时得到。最大熵法是比较精确且自动选择阈值的方法。现摘录如下<sup>[5]</sup>——

设一幅灰度图像的灰级数为  $L(1, 2, 3, \dots, L)$ ， $f_i$  为灰度为  $i$  的像素数，图像的总像素为：

$$N = \sum_{i=1}^L f_i$$

将直方图表示为概率密度分布：（将  $P_i$  记为概率密度）

$$P_i = \frac{f_i}{N}, \quad \sum_{i=1}^L P_i = 1, \quad P_i \geq 0$$

假设阈值  $S$  将图像分为两类（物体和背景），即  $C_0 = \{1, 2, 3, \dots, S\}$ ， $C_1 = \{S+1, S+2, S+3, \dots, S+L\}$ ，则  $C_0$ 、 $C_1$  的信息熵  $H_0(S)$ 、 $H_1(S)$

将  $H_0(S)$  定义为：

$$H_0(S) = - \sum_{i=1}^S \frac{P_i}{P_0(S)} \ln \left[ \frac{P_i}{P_0(S)} \right] = - \frac{1}{P_0(S)} \sum_{i=1}^S P_i [\ln P_i - \ln P_0(S)] = \ln P_0(S) + \frac{H(S)}{P_0(S)}$$

其中，

$$P_0(S) = \sum_{i=1}^S P_i \quad H(S) = - \sum_{i=1}^S P_i \ln P_i$$

将  $H_1(S)$  定义为：

$$H_1(S) = - \sum_{i=S+1}^L \frac{1}{P_1(S)} \ln \left[ \frac{P_i}{P_1(S)} \right] = \ln[1 - P_0(S)] + \frac{H(L) - H(S)}{1 - P_0(S)}$$

其中，

$$P_1(S) + P_0(S) = 1 \quad H(L) = - \sum_{i=1}^L P_i \ln P_i$$

则对应的最大熵测度准则  $\varphi(S)$  为：

$$\varphi(S) = H_0(S) + H_1(S) = \ln\{P_0(S)[1 - P_0(S)]\} + \frac{H(L) - H(S)}{1 - P_0(S)} \quad (4)$$

最佳阈值  $S'$  满足：

$$\varphi(S') = \max_{1 \leq S \leq L} \{\varphi(S)\} \quad (5)$$

阈值法分割原理即为通过设定不同局部的不同阈值将图像分为若干类。由上文引用的算法可以求得用户面部图像的最佳阈值，通过与“标准阈值”的对比，即可将图片分为不同的部分，包括——眼角、嘴角、眼睑等等。





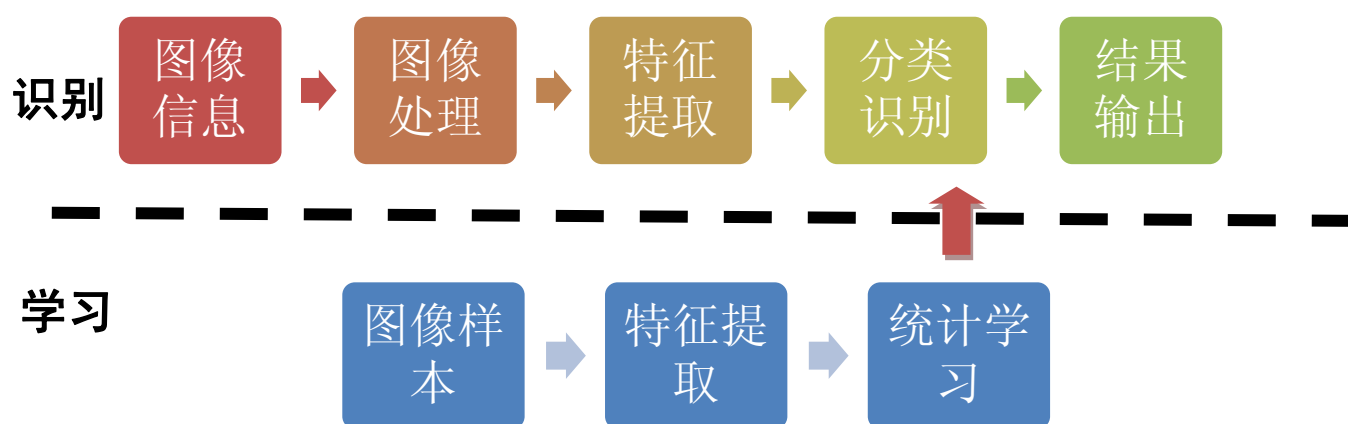
空间转换

边缘提取

### 3.3.2 图像识别

传统的图像识别方法一般可以归为五大类，即模板匹配法、统计识别法、结构识别法、模糊识别法和人工神经网络识别法。经过参考相关资料<sup>[6][7]</sup>，本文采用了统计模式识别法。简述如下：

统计模式图像识别法的步骤如下图所示：



输入的数字图像信息是计算机可识别的电子信号，如 JPG 图像或者 BMP 图像；其目的就是对样本图像进行滤波、恢复其中便于从原始信号中的能够反映图像特征的信息（在本文中如系统预置的“黑眼圈”、“向下/上嘴角”、“眼角”等特征信息，便于与采集的图像进行匹配）；最后一个步骤是分类识别、输出结果。在确定识别判别函数时需要大量的多种类别的样本，通过对其特征的计算统计从而设定特征向量和相应的参数，也就是下图中所示的训练学习过程。

由上文可知，若用户面部图像经过处理已经提取了  $N$  个特征，而样本集分为  $m$  类，那么识别图像的模式可以认为是  $N$  维向量空间的向量  $X$ ，即：

$$X = [X_1, X_2, X_3, \dots, X_N]^T$$

模式的类别可设为： $\omega_1, \omega_2, \dots, \omega_m$  我们最终识别的内容就是要判断上述的“ $X$ ”属于“ $m$ ”中的那一种类。然而这种模式识别仍然有一个重要的关键问题，即判别函数的选取会影响到识别的精度，尚有待进一步研究。



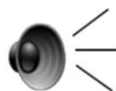
### 3.4 语音控制识别技术

本节的灵感来源于2011年与iphone 4S一同发布的Siri应用。区别于传统的语音控制识别技术，如Voice Actions，Siri不仅提供了非常坚实可靠的声音识别引擎且高识别度令人称奇，同时还可以识别语法结构“不那么严谨”的指令，凸显其与用户的“心有灵犀”，而这正是作者所期望的。

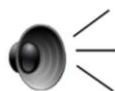


Siri的主要功能听起来相当不一般，使用者可以通过声控、文字输入的方式，来搜寻餐厅、电影院等生活信息，同时甚至是直接订位、订票；另外它的适地性服务的能力也相当强悍，能够依据用户默认的居家地址或是所在位置来判断、过滤搜寻的结果。

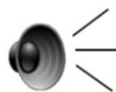
作者的思路即把Siri或者与Siri类似的功能移植到本“问候系统”，同时需要增加部分Siri所不具备的功能——既然是智能问候系统，且通过图像分割识别我们已得到了用户的健康状况、精神面貌，那么系统就应该具有“针对用户健康状况及精神面貌提出建议”的功能，比如：



“您没休息好吧，记得别熬夜，要早点睡哦！”



“高兴点！我给你讲一个笑话！”



“每天要喝够八杯水哦！”

.....

系统的语音库还需定期自动联网更新，以避免日复一日的单一回复引起用户反感。

此外，对于用户提出的问题，如“附近堵不堵车？”“我要带伞吗？”“去开会穿这身西装行不行？”本系统能够根据用户的日常喜好、性格提出针对性建议，即使不能解决，也会联网查找。此外，除了语音反馈，本系统还可根据用户需求将所查找到的数据（如离线地图、路况、天气预报）通过邮件发送至用户电

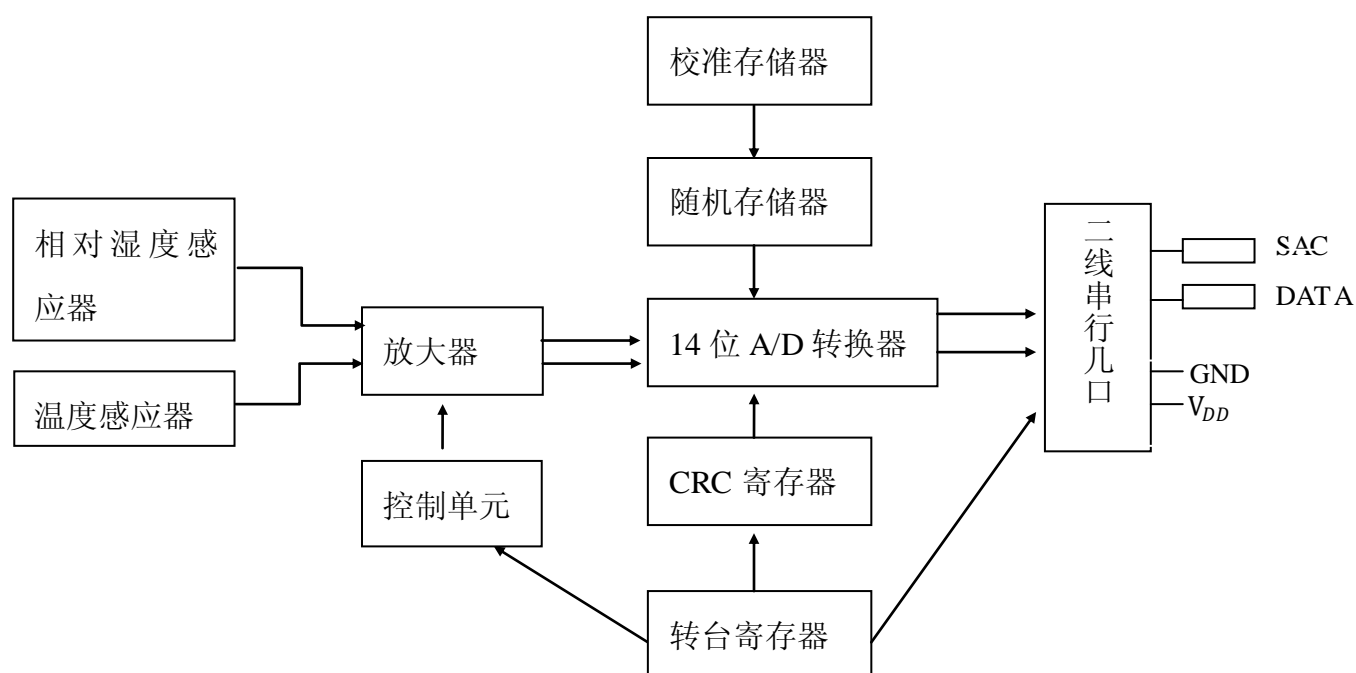
脑或手机（前提是需要预设），是用户的使用变得更加高效、灵活。通过互动镜面以及互联网平台，用户与本系统可以实现积极的交互。

语音控制识别技术当下已较为成熟，将 Siri 或与其类似的技术实现移植，稍稍经过改进便可运用于本系统，实现系统对用户“主动”的问候，从而让人们的居家生活变得更为人性化且智能化。

### 3.5 温湿度感应技术

由于本系统主要应用于浴室、卫生间，所以必须考虑到防水防雾的要求。本系统在启动后将采用温湿度传感器对周围环境的温度及湿度进行实时监控，当湿度过高（ $\geq 80\%$ ）时进入休眠状态保护内部元件；当环境温度过高（ $\geq 45^\circ\text{C}$ ）时还能提醒用户及时通风换气，以达到“智能”的要求。

通过查找资料，本文决定采用 SHT11 传感器<sup>[8]</sup>。简述如下：SHT11 传感器默认的测量温度和相对湿度的分辨率分别为 14 位、12 位，通过状态寄存器可降至 12 位、8 位。湿度测量范围是  $0\sim 100\%\text{RH}$ ，对于 12 位的分辨率为  $0.03\%\text{RH}$ ，测温范围为  $-40^\circ\text{C}\sim +123.18^\circ\text{C}$ ，对于 14 位的分辨率为  $0.01^\circ\text{C}$ 。每个传感器芯片都在极为精确的湿度室中进行标定。校准系数以程序形式储存在 OTP 内存中，在测量过程中可对相对湿度自动校准，使 SHT11 具有 100% 的互换性。其内部结构框图大致如下：



通过温湿度传感器获得的环境温/湿度数据，经过处理，本系统将能够判断房间环境是否适合用户使用。与前文所述的语音识别控制系统相结合，当卫生间过于闷热潮湿时，系统则将提醒用户“开窗通风”、“请缩短洗漱时间”，若能与浴室的排风扇相连，这样一来，就算长时间无人应答，本系统也可以自动对房间进行换气，无疑用户的人身安全又有了较大保障。

在此基础上，还可以将“CO 浓度检测”等功能作为本系统的后续研究方向，使其功能更加全面、人性化且安全可靠。

## 第四章 结束语

本文的创意其实也是受到“家居智能化”的启发，旨在使传统居家物品能够变得“智能”，善解人意。

本系统的主要功能是“问候”，即在用户照镜子的同时可以对其健康状况、精神面貌进行简单评估并给出相应的建议；与此同时，强大的语音控制识别技术将使用户能与本系统轻松交互，即系统可以针对用户的询问如天气、路况等做出相应回答；除了需指纹扫描启动外，用户的整个操作过程都通过语音完成，使用简便，充分考虑到了用户洗漱时操控的不方便，彰显其设计人性化。

重要的是，本系统所涉及到的技术，包括指纹识别、语音控制识别、图像分割识别、温湿度感应技术，都是目前发展相对成熟、却又有很大提升空间的。故本系统的可行性极高，且其进一步提升空间可谓宽广（如文中提到增加 CO 含量检测等等），各项技术的进一步发展也有望解决研发成本问题。智能问候系统在家具行业目前尚不常见，所以作者认为，本系统的发展潜力巨大，且它的提出，或多或少将会对现代家居产业产生一定影响，且这种影响是正面的、符合未来“家居智能化”这一大方向的。

限于专业知识的不足，论文可能缺乏说服力。部分技术的采用或许会有更好的改进方案。敬请专家及读者给予指正，谢谢！

### 参考文献

- [1] 田捷, 陈新建, 张阳阳, 杨鑫, 何余良, 李亮, 谢卫华, 郑志鹏•指纹识别技术的新发展•自然科学进展•2006 年 4 月第 16 卷第 4 期
- [2] 刘璇, 李建平•人脸定位与跟踪技术的研究与实现•电子科技大学•中国知网 2011 年第 3 期
- [3] 刘庆祥, 蒋天发•彩色与灰度图像间转换算法的研究•武汉理工大学学报(交通科学与工程版)•2003 年 6 月第 27 卷第 3 期
- [4] 林开颜, 吴军辉, 徐立鸿•彩色图像分割方法综述•中国图像图形学报•2005 年 1 月第 10 卷第 1 期
- [5] 张家树•灰度图像的最大最小优化熵阈值分割•电讯技术•1995 年 12 月第 35 卷第 6 期
- [6] 吴冬夏, 蒲晓蓉, 陈莲娜•基于小波分解和颜色信息熵的浮游生物图像识别技术研究•电子科技大学•中国知网 2012 年第 2 期
- [7] 卢力, 田金文, 柳健•统计模式识别研究进展•军/民两用技术与产品•2003 年 11 月•文章编号: 1009-8119(2003)11-0039-04
- [8] 冯显英, 葛荣雨•基于数字温湿度传感器 SHT11 的温湿度测控系统•PROCESS AUTOMATION INSTRUMENTATION Vol. 27 No. 1 January 2006 •©1994-2010 China Academic Journal Electronic Publishing House. <http://www.cnki.net/>