人脸识别解锁手机的奥秘

1.引言：

随着人脸识别技术的快速发展，手机的人脸解锁凭借其准确有效的解锁效果已经成为最广泛使用的生物识别技术之一。早在 2008 年，著名华人学者汤晓鸥就发明了准确度极高的人脸识别算法并且荣获 cvpr2009 最佳论文，在 2017 年国内诞生了第一款人脸解锁手机金立 s10，而如今在国内如旷视科技，商汤科技也已经成为全球脸部识别算法中的佼佼者，可以说中国是人脸识别领域的领跑者之一。而在这篇文章中，我想通过对人脸识别以及手机摄像头提取图像的介绍来科普一下手机人脸解锁中的奥秘。

2.论文主体：

整个人脸解锁环节由人脸识别部分和摄像头提取图像信息部分构成。而人脸识别的整个检测环节主要由人脸检测和人脸识别两个环节组成。

第一个环节人脸检测较为简单，其目的是刨除一切非人脸部分的干扰，想象一下，如果识别的时候摄像头不仅把人脸，也把周边的环境作为参数传入 f（x）识别函数，那必然得不到准确的结果。人脸检测的技术要求相对低一些，只需要知道有没有人脸以及人脸在照片中的大致位置即可，在网络上有较多开源代码可以帮助开发者实现人脸检测的功能。

而第二个人脸识别的环节原理就较为复杂，其复杂体现在如何去构造一个准确率较高的 f(x)识别函数去形成特征值。在了解特征值是如何形成之前，我们先要了解特征值的本质是什么。实际上特征值并不是一个数字，仅仅一个数学无法表现出“特征”两个字，通常我们用多个数值组成的向量表示特征值，向量的维度即其中的数值个数。而特征值的维度越大越好吗？显然不是，从 Google 的研究结果表面，128 个数值组成的特征向量结果最好。

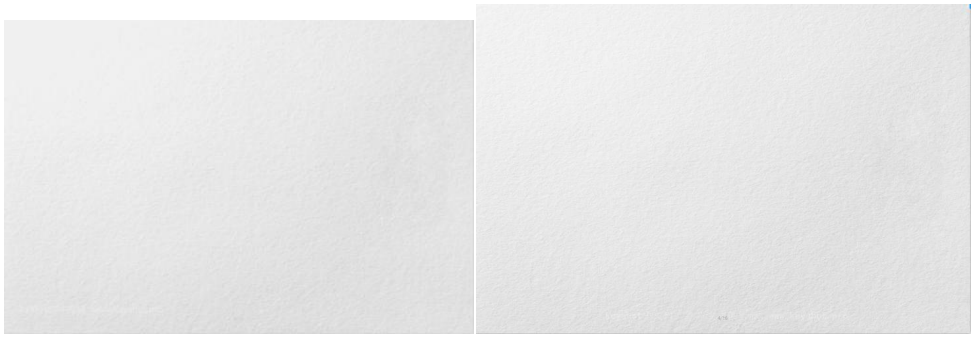
Table 4. **Image Quality.** The table on the left shows the effect on the validation rate at 10E-3 precision with varying JPEG quality. The one on the right shows how the image size in pixels effects the validation rate at 10E-3 precision. This experiment was done with NN1 on the first split of our test hold-out dataset.

#dims	VAL
64	86.8% ± 1.7
128	87.9% ± 1.9
256	87.7% ± 1.9
512	85.6% ± 2.0

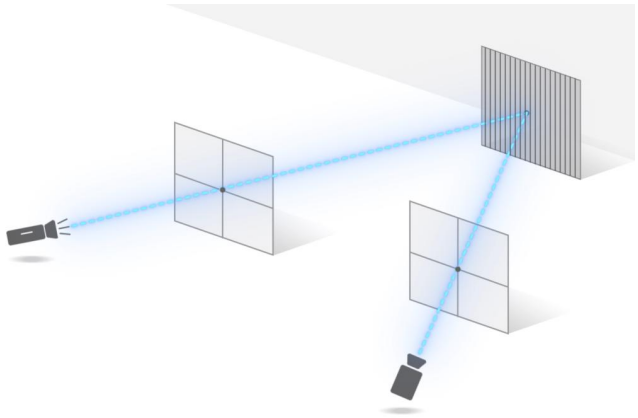
（图来自于论文 https://arxiv.org/abs/1503.03832）

在求特征值的过程中，算法要处理的数据是巨量的，假定给出的人脸照片是 100*100 像素大小，由于每个像素有 RGB 三个通道，每个像素通道由 0-255 范围的字节表示，则共有 3 个 100*100 的矩阵计 3 万个字节作为输入数据。很显然，如果人工对函数进行修正是不切实际，所以在人脸识别技术在 60 年代被提出一直到 21 世纪初都处于不温不火的状态，直到深度学习理论的爆发，人脸识别才有今天这样磅礴的发展。在如今人脸识别开发中，开发者会给一个初始的 f（x）函数，这个函数有这大量参数而且结果并不准确，开发者的目的是对于任何输入的照片 x，通过函数计算以后都能得到正确的特征值 y，所以此时对该函数进行大量人脸识别的测试，通过深度学习结果的反馈反过来不断自行修正函数中的参数来让下一次结果逼近于正确的特征值 y，最终得到较为准备的识别函数。

有了算法，没有硬件的更新来对图像进行准确的输入，那人脸识别就像纸上谈兵。在人脸解锁刚推出时，很多人质疑：“万一有人拿出一个本人照片来，机器会不会被骗过去？”显然是不会的，因为真人有轮廓，而照片没有。传统的摄像头，是通过视觉差来测量轮廓，但缺点很明显对于一些光滑、或者缺乏纹理的表面，摄像头难以分辨距离。在这种情况下，为人脸识别而生的结构光就排上用途了，在结构光中有两个摄像头其中一颗摄像头，变成一个投射器，主动往目标上投射纹理，再由另一颗摄像头捕获，就可以完成对人的脸部轮廓的构建。这就是最基本的结构光模型了。



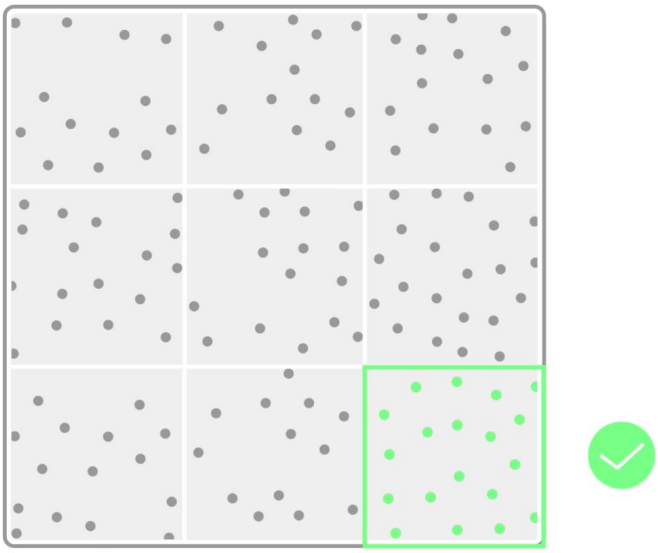
（同样一面墙，你能分辨哪个离你更远吗？）



（结构光）

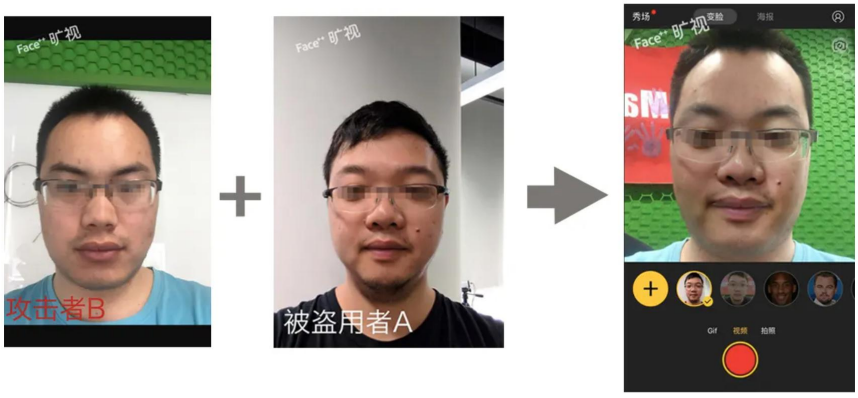
在对人脸图像进行提取时，结构光通过点阵投影来构建人脸图像，再将数据作为参数传给算法，在这个过程中会以伪随机的形式，加入一些空帧，比如亮一帧、停一帧，再亮一帧；或者先空两帧，再亮一帧。每次解锁，点阵亮起的时间间隔，都不一样。在这里也是利用了哈希加密的原理。就像是一个特工在和总部联络：我们相约，在星期一、星期三、和星期五，分别给你三份情报。那如果总部在星期二收到了一份情报，哪怕情报信息完全正确，伪造得天衣无缝，但只要时间因为每次交接的时间，都是伪随机的，具体哪一天来交接，只有特工，和总部，两个人知道。这就可以防止有心人士，拿伪造的红外图来欺骗系统。即使传进来的图像是正确的，系统依旧不会对其进行识别。

在前面我提到过，人脸识别所需的像素数量是惊人的，仅仅一个手机 cpu 想快速的完成检测是很难完成的，而用户也不会接受在手机旁等个十秒来解锁手机，如果单纯为了安全性，那就直接输入密码关闭人脸解锁了。所以实际的人脸解锁里运用了抽样检测的原理，手机将人脸图像分隔成一个个小块，再将一个个小块作为单位进行匹配，图像快越小进行匹配时要进行的迭代次数就更多，所以速度会慢下来。此时厂家就可以根据自家手机芯片的运算能力来对图片分块大小进行限制，从而控制合理的解锁时间。



3.结论：

人脸识别技术给人类社会带来了巨大的经济收益，但我们也需要知道，罗马不是一天建成的，正是科学家在深度学习，在传感器摄像头以及密码学方面的一次次突破，才有了今天的脸部解锁。但我们也需要知道，目前的人脸识别仍有不足，在 1：N 模式下很容易出现错误，很多人都经历过自己去小区或者校园门口扫脸部门禁，结果扫出一个从来不认识的人这种尴尬。在安全性上也有一些漏洞，这仍需从业者们的努力去完善。但我相信人脸识别的运用前景是十分广泛的。



（攻击者通过人脸融合来完成验证）

4. 参考文献

[1]iotcaf.人脸识别技术及应用概览
[N/OL].httttps://zhuanlan.zhihu.com/p/36309038 ，2019-05-22：

[2]zealer.iPhone 13 的刘海，为什么缩小了？
[N/OL].https://mp.weixin.qq.com/s/9ZwCO7WGYeiPbXUlYAH5iQ ，2022 年 1 月 26 日;