# Snapchat的过滤器：计算机视觉如何识别你的脸

在那些无聊的时刻，当你在玩“过滤网”的时候，伸出舌头，啃食你的容貌，并想办法让花冠正好贴在你的头上，你肯定有过这样的时刻：在技术层面上，你想知道发生了什么事，Snapchat是如何与你的脸匹配的去看动画？

经过两个星期的网上调查，我很感激终于在幕后看到了一眼。结果表明，该产品是计算机视觉应用的一个实例，是各种人脸识别软件背后的主要燃料。

#### 技术

这项技术来自乌克兰一家初创公司Looksery，这是一款允许用户在视频聊天和拍照时修改面部特征的应用程序。Snapchat于2015年9月以1.5亿美元的价格收购了这家基于敖德萨的变脸创业公司。据报道，这是乌克兰历史上最大的科技收购案。

他们开拓了计算机视觉的广阔和迅速发展的领域。计算机视觉可以被认为是计算机图形学的直接反面。计算机图形学试图从三维模型生成图像模型，而计算机视觉则试图从图像数据创建三维空间。计算机视觉在社会中的应用越来越广泛。这是你扫描支票的方式，数据是从行里提取出来的。你可以用手机存支票。这就是Facebook如何知道你照片中的人，自动驾驶汽车如何避免撞人，以及你如何给自己一个狡猾的鼻子。

#### Snapchat过滤器的工作原理

Looksery对他们的工程技术保密，但每个人都可以在线访问他们的专利。Snapchat过滤器使用的计算机视觉的特定区域称为图像处理。图像处理是通过对所提供图片上的每个单独像素执行数学运算来对图像进行变换。

1 - 面部检测：

第一步工作如下：给定一个输入图像或视频帧，找出所有当前的人脸并输出其边界框（即矩形坐标的形式：X，Y，Width&Height）。

自21世纪初以来，人脸检测一直是一个亟待解决的问题，但也面临着一些挑战，包括检测和人脸。最广泛使用的技术是方向梯度直方图（简而言之）和支持向量机（Support Vector Machine）的组合，在获得高质量图像的情况下，可以获得中等到相对较好的检测率，但这种方法至少不能在CPU上实现实时检测。



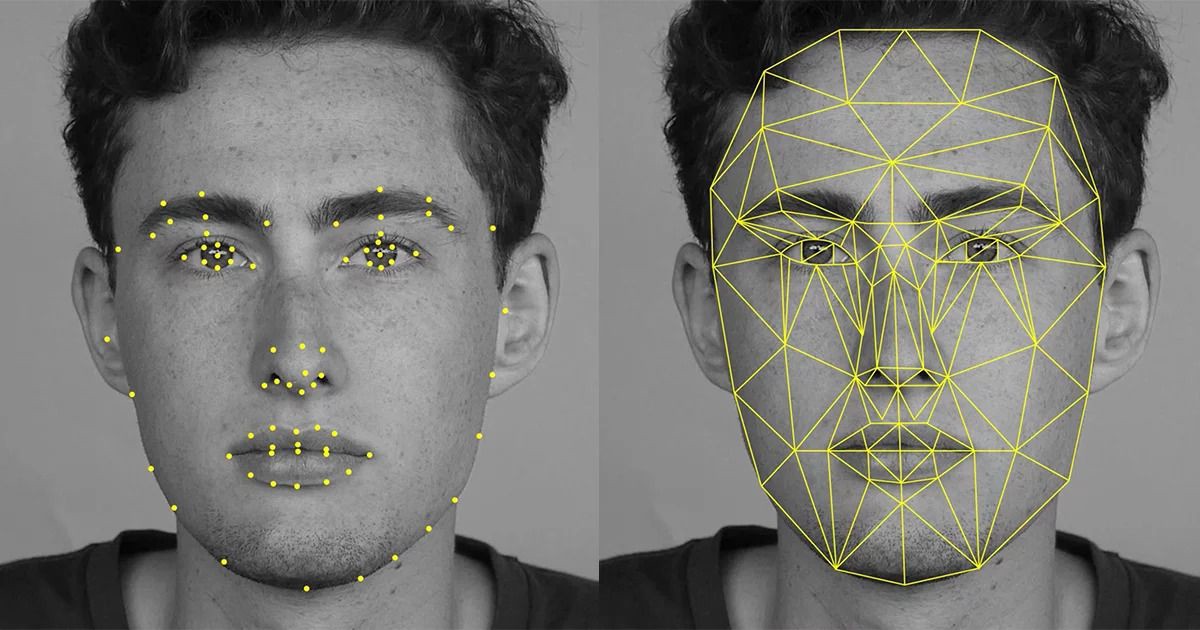
以下是HOG/SVM检测器的工作原理：

给定一个输入图像，计算该图像的表示，该图像是原始图像的多比例缩小版本的金字塔。对于棱锥体上的每个条目，都使用滑动窗口方法。滑动窗口的概念非常简单。通过在具有恒定步长的图像上循环，以不同的比例提取通常大小为64 x 128像素的小图像块。对于每个面片，算法会决定是否包含一个面。计算当前窗口的HOG，并将其传递给支持向量机分类器（线性或非线性）以进行决策（即是否是人脸）。当使用金字塔时，通常会进行（简称NMS）操作以丢弃堆叠的矩形。您可以阅读更多关于HOG/SVM组合的信息。

2 - 面部标志：

这是我们分析阶段的下一步，工作如下：对于每个检测到的人脸，输出该人脸的每个成员或面部特征的局部区域坐标。这包括眼睛、骨头、嘴唇、鼻子、嘴……通常以点（X，Y）的形式表示的坐标。

对于给定一个边界框（即目标人脸的裁剪图像）的CPU来说，提取面部地标是一个相对便宜的操作，但是对于程序员来说，除非使用一些不太快的机器学习技术，例如训练和运行分类器，否则很难实现。



你可以找到更多关于提取面部地标的信息或此PDF:。在某些明显有用的情况下，将人脸检测和标记提取结合到一个操作中。

3 - 图像处理

现在已经检测到人脸，Snapchat可以使用图像处理将特征应用到完整的人脸上。然而，他们选择更进一步，他们想找到你的面部特征。这是在主动形状模型的帮助下完成的。

主动形状模型是通过在数百到数千幅图像上手动标记面部特征边界来训练的面部模型。通过机器学习，创建一个“平均脸”，并将其与提供的图像对齐。当然，这个普通的脸并不完全适合用户的脸（我们都有不同的脸），所以在拟合完这个脸之后，会检查“普通脸”边缘周围的像素，以寻找阴影的差异。由于该算法经过了训练（机器学习过程），它有一个关于某些面部特征应该如何显示的基本框架，因此它在给定的图像中寻找相似的模式。即使一些最初的改变是错误的，通过考虑它已经修正的其他点的位置，算法也会修正它所做的关于它认为你脸的某些方面在哪里的错误。然后，模型将调整并创建网格（一个可以随面移动和缩放的三维模型）。



整个面部/特征识别过程是在你选择过滤器之前看到白网时完成的。然后，过滤器通过增强或在其上添加一些东西来扭曲提供的面部的某些区域。

#### 从过滤器到面交换

几个月前更新的Snapchat版本具有与朋友交换面孔的功能，无论是实时交换，还是从图库访问一些面孔。请注意面形状是如何可见的，这就是统计模型所在的位置。它可以帮助Snapchat快速将您和朋友的脸对齐并交换功能。

定位完所有特征后，应用程序将沿面创建一个网格，逐帧粘贴到每个点。现在可以根据Snapchat的感觉编辑和修改此网格。



有些镜片会让你扬起眉毛，或者张开嘴。这也相当简单，但它需要更多的算法来暗示。

* 相对而言，口腔内部比较暗。这就泄露了你的嘴。
* 当它发现用户已经扬起眉毛时，眉毛相对于其他面部特征的变化会被考虑在内。

如前所述，这项技术并不新鲜。但要在移动平台上实时执行所有这些过程，需要大量的处理能力和一些复杂的算法。这就是为什么Snapchat认为最好花1.5亿美元收购Looksery，而不是仅仅建立它的平台。

#### 结论

我希望这是一个信息丰富的和挠痒痒的好奇心，就像它做了我的。现在，我将更深入地探索Snapchat过滤器，测试我最喜欢的面部镜头，了解并欣赏幕后发生的所有计算机视觉。

— —

如果你喜欢这首曲子，我喜欢你按一下拍手键，这样别人可能会偶然发现。你可以在上找到我自己的代码，以及我的更多写作和项目。你也可以跟着我，或者。

其他资源：

* （技术字节）
* （技术百）
* （语音）

#### 这篇文章发表在Startup上，这是媒体最大的创业出版物，其次是294522+人。

#### Subscribe to receive our top stories here.