

硕 士 研 究 生 读 书 报 告



题目自动人脸图像处理技术读书报告

作者姓名 金乔觉

作者学号 21651070

指导教师 李启雷

学科专业 移动互联网与游戏开发

所在学院 软件学院

提交日期 二○一七 年四 月

摘要

自动人脸图像处理技术，是一种新兴的图像处理技术。本文参考的文献提出了一种新的框架，自动编辑照片中眼睛部分，利用用户个人照片收集找到一个“好”的参考眼睛和转移到目标图像，开发了一个综合管线三维面估计，图像变形，图像，图像融合，图像合成的自动分割，并以实现高度可信的结果。

**关键词**：人脸图像，自动编辑，眼睛

Abstract

Automatic face image processing technology is a new image processing technology. This paper reference literature puts forward a new framework of automatic photo editing in eye part, using personal photo collection to find a "good" reference eyes and transferred to the target image, developed a comprehensive 3D pipeline estimation, image morphing, image segmentation, image fusion, image synthesis, and in order to achieve highly reliable results.

**Keywords：**Face image, auto edit, eyes

1. 技术简要介绍

在摄影和分享照片的过程中，人的面孔通常是观察者十分感兴趣的内容。然而，拍摄好的肖像具有一定的难度，具体原因：（1）.拍摄主体变得紧张时，面对镜头可能露出不自然的表情。（2）.闪光会导致拍摄主体闭眼。（3）.角膜反射或瞬目反射。这些原因，导致在实际拍摄的过程中，拍摄主体很难在整个过程中一直保持睁大眼睛的状态，很难保持良好的拍摄效果。在最近的一些研究中，一些研究人员试图利用相关知识预测效果比较好的面部表情，并设计一个应用程序来帮助用户户外训练最吸引人的表情。然而，还有一个更好的方法就是直接应用图像编辑来改变照片中的面部表情。因为在实际拍摄过程中，很难确保每个人都在最完美的时候按下快门，所以这种图像编辑照片技术，是非常实用的。但是，一般的图像编辑工具，如Abode, PS图像处理软件，需要大量的相关专业知识。在这篇报告所参考的文章中，作者介绍了一种处理照片中眼睛部分的完全自动的方法，以提高照片的质量，一个全自动的眼睛编辑方法，不需要任何互动。如果需要的话，还能提供给用户几个用于参考的图像。

作者提供的方法，实用过程如下：

1. 给定一个闭眼的图像
2. 选定几个合适的图像
3. 筛选图像中的眼睛区域
4. 替换原始图像中的眼睛区域

除了改变眼睛部位，此系统还可以用来修正目光来提高拍摄的质量和效果。

1. 相关工作

整个过程，设计的有面部编辑，局部颜色转移，图像合成，人群反馈。

1. 脸部图像编辑

几十年来，人脸一直是图像增强和编辑的主要目标。Blanz和Vetter提出了一种三维形变模型合成，从一个单一的到新的姿态，表情，或者是照明。Nguyen等人应用子空间的方法将眼睛和胡须摘除。Leyvand等人开发了一个脸部变形系统提高拍摄效果。Kemelmacher等人利用数年的人的数据合成一个人不同年龄的面孔。一个更接近的研究是一个基于实例的方法，其中某些属性的例子被转移到输入目标中。2008年，Bitouk等人开发了一个能自动将三个整脸从一个图像自动转移到另一个图像的系统，这个系统后来被扩展为全脸性更换视频。2009年，Guo等人介绍了一种传递面部化妆法，2010年，Joshi等人开始采用个人实例改善人脸图像质量。2014年，有人建议将图像风格由专业转移到休闲。2011年，有人开发了一个基于实例的交互式表情编辑系统。不同于之前的所有系统，本系统中作者着眼于现场的眼睛外观，同事也考虑到颜色和纹理，他们结合相关的实例和人的反馈寻求达到最好的结果，另外，这个系统还是全自动的。

1. 图像合成

在复制粘贴编辑，合成方法获得高品质的结果是至关重要的。常用的技术包括alpha混合，多尺度合成，梯度域混合。内容的无缝克隆通常通过梯度域的合成来实现。这些方法在遇到文物时，通常会碰到前景和北京的边界不匹配等问题。2013年，Tao等人建议“隐藏”纹理区域的错误，以避免明显的出血文物。2010年，有人开发了一种多尺度直方图匹配的方法，它允许纹理自然混合。在眼睛编辑的背景下，彩色伪影非常引人注目。因此，原作者结合阿尔法混合和无缝克隆的优点，开发了一个简单的技术，以达到合理的结果。

1. 逆向绘制

逆渲染技术的最新进展实现了在视频里动态重建面部几何，并实现了一些面部操作，而在本系统中，作者只需要关注一个编辑的眼睛，它是基于图像输入，而不是非常精确的几何面。

1. 通过在线群体学习

在Shih和Bitouk等人的实验中，例子是由有经验的手工绘制人员选择的。这些方法并没有根据人类的喜好进行优化。就像计算机视觉的最新进展成果，作者在考虑了大量人类喜好后，寻求通过向在线群体学习的方法来改善系统。人类标记的数据已被用于许多不同的任务。举个例子，2014年，Kiapour等人通过向人群采集一定的判断风格，研究揭示个人风格的服装。拉丰等人提出了一种高层次的图像编辑方法，使用户能够收集群众来源的注释来调整场景的属性。还有人开发了一个通过向人群评价学习，来提高肖像吸引力的系统。

1. 自动开眼

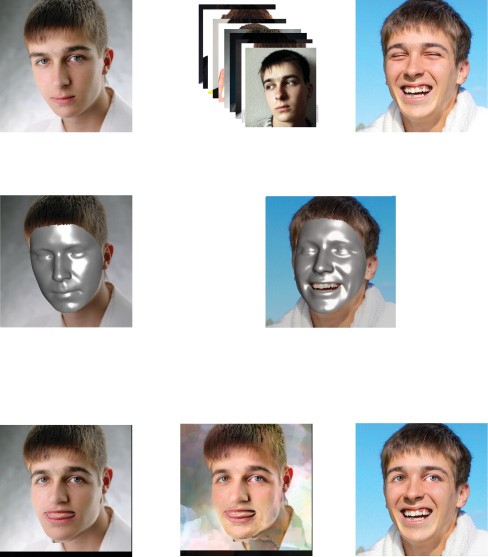


图1,

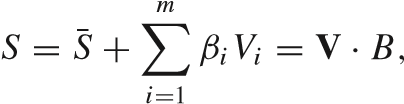
图1是整个眼睛编辑管道的工作流程，给定一个图像与修正图像，这个系统通过另一个输入例子取代这个图像中的眼睛。首先进行三维人脸模型的拟合。这里使用到一个人脸检测器和一个拷贝混合图像合成方法。另外本地对比度和照明也是非常重要的参考因素，通过对齐，图像协调，扭曲参考面匹配目标面，最后经过局部彩色矫正，得到最终结果。鉴于需要编辑眼睛的目标图像，以及具有所需眼睛外观的参考，系统通过以下步骤自动传输眼睛外观：（1）3D人脸拟合（2）姿态校正（3）局部颜色调整（4）图像混合。

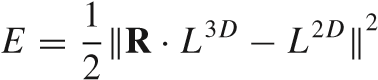
（1）3D人脸拟合

让n个3D点的串联表示三维面：

S =（X1，Y1，Z1，…，xn，yn，Zn）

所有的3D空间的面孔通过三维形变模型。进行主成分分析对三维人脸形状的数据集获得Eigenshapes（特征向量）的面空间用VN×M = [ v1，…，VM ]（只有第一个明显特征向量的选择）。一个三维人脸可以近似由一个线性组合表示如下：



¯S表示平均形状在三维人脸数据集。在一个图像中，二维人脸的形状是三维人脸形状投影到图像平面所产生的结果：S2D = R•S3D。可以通过最小化3D坐标和2D坐标图像检测之间的误差，来恢复投影矩阵R和B的形状特征： 

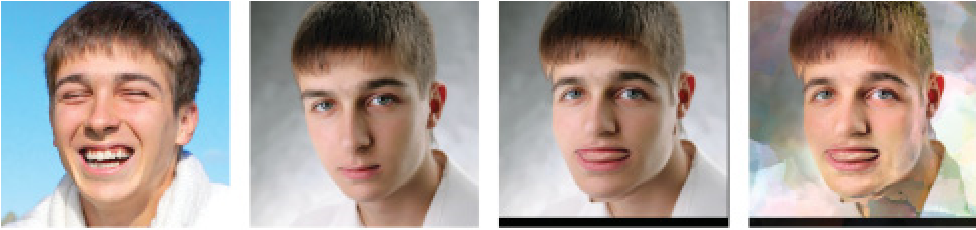


图2

图2采用多维直方图匹配进行局部颜色调整。比较从图b脸的颜色分布（扭曲和对齐参考）到目标面部颜色分布得到图d。请注意，我们只使用眼部区域图c合成为目标。

每一次迭代中的优化，为三维人脸姿态的变化，沿封闭轮廓的二维人脸坐标对应的投影的三维人脸不同顶点。为了处理这种情况，我们使用杨等人提出的图像拟合算法。为了优化R和B，因为作者这个系统专注于眼睛，他们在眉毛，眼睛和鼻子施加更高的权重（正常体重的4倍）的地标。这减少了由于表达变化的拟合误差，但强劲捕获身份和姿态变化。

（2）姿势

对于图像变形，作者采用杨等人的三维点位移管道。如前文所述，拟合三维形变人脸模型参考和目标建立明确的三维顶点到顶点的两个图像之间的对应关系。通过将图像投射到其对应的三维人脸形状相同的拓扑结构，建立了一个像素的二维位移场与器件像元强劲捕捉姿势旋转的非线性效应。利用二维位移场经参考图像大致对准目标眼睛。

（3）局部色彩调整

眼睛粗对准到匹配目标后，第二步就是协调图像匹配的整体肤色，照明，从目标图像的局部对比度等方面。在本节中，我们介绍了一种新的方法，多维直方图匹配，以实现强大的局部颜色转移：

给出两组n维数据Xn ，Yn，以及他们在n维空间H分布h（X）和h（Y），寻求一种映射函数fn→N（•）：

Zn×M1 = f（x）

这样的Z的分布相匹配的Y（H）= H（y）的分布。这是n维空间中的直方图匹配问题。当n = 1时，直方图匹配的问题，可以很容易地解决一个离散的查找表。

2005年，Pitie等人应用多维直方图匹配ING颜色传递当n = 3（RGB颜色空间）。对于局部颜色传输的应用，作者采用多维分布匹配的算法，但额外的尺寸会加强空间局部性匹配。

像RGB直方图的图像统计是全球性的，不捕捉当地的影响，如照明和阴影变化的脸。为了模拟局部效应，而颜色匹配，我们建议重写图像表示从I (3)(x,y) = [r(x,y),g(x,y),b(x,y)]到I (5)(i) = [ri,gi,bi,xi,yi]。

I是像素指数。这种表示显式地编码了图像像素的局部性。因此，可以建立一五维直方图的图像编码的局部性和语义像素信息。然后进行多维直方图匹配目标。

由于表示是独一无二的，从*ref*W精确匹配为目标的结果将是目标本身。作者将去除所有图像坐标中x和y的变化（第三、四维度的I（5））来保持*ref*W纹理。由于人们只对眼睛周围的区域感兴趣，所以匹配函数每次只计算迭代中的人脸区域。在实践中，会平滑地处理匹配功能，以避免量化的彩色文物。

（4）BFGP合成

应用局部色彩传递到refw之后，得到了一个形象的例子refwlc预处理。剩余的眼编辑步骤是对refwlc眼部粘贴到目标。在最后的合成，作者使用密集光流重新调整眼部的影响及眼部区域的目标。这消除了大部分的差异所造成的翘曲。差异是由表达式的变化或不准确的姿态估计造成的。然而，作者不使用光流扭曲图像。相反，计算光流的平均运动，并把它应用到refwlc。随后，在target和refwlc，使用眼睛区域的地标和应用图形切割对眼部区域边界进行自动定义。

优化梯度域混合区的概念类似于Agarwala等人所做的工作。作者定义区域上的图像梯度域。具体而言，周围的感兴趣区域（ROI）的高梯度区域更应该成为被切割的一部分。在作者的应用中，ROI是眼睛区域定义的眼睛地标。在以前的复制粘贴为基础的方法中，合成得到的图像梯度域使用泊松混合。梯度域合成方法产生的图像无缝复合材料只有在图像对边界条件大致匹配。现实生活中的照片是在不受限制的光照条件下拍摄的，局部颜色转移算法相匹配的照明和局部阴影的目标和参考规模较大。

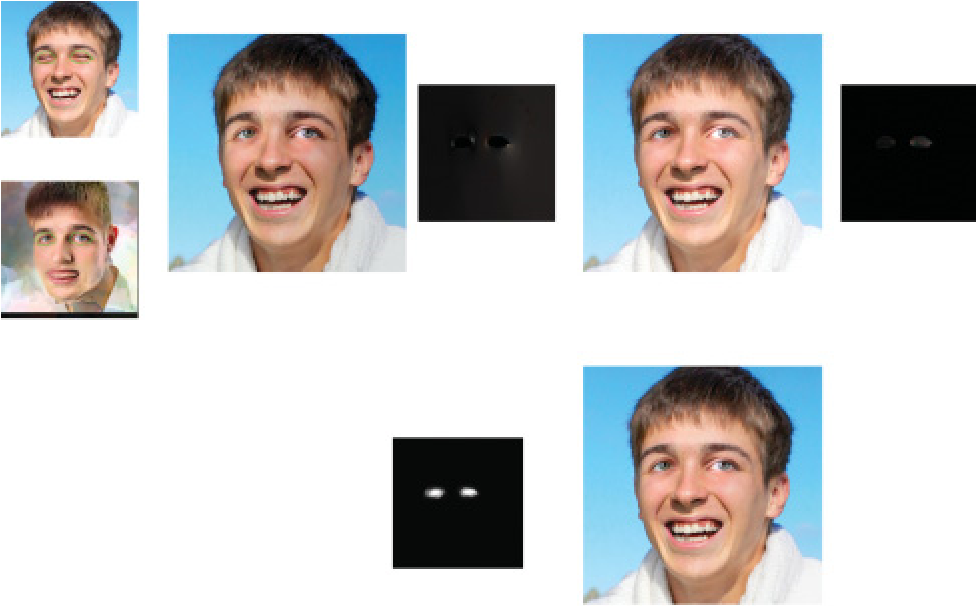


图3

图3。BFGP合成：结合梯度域融合和alpha混合。执行梯度域融合两次， refwlc.gdbf是梯度域拼接的结果（即眼睛）被保留了下来，在gdbb背景保存。对于gdbf，误差是背景的平均像素值误差。对于gdbb，误差是前景平均像素值误差。通过图边切割优化混合边界。

因此，边界条件不匹配可能还是由于小规模的局部阴影和阴影眼睛周围存在。在不匹配的边界条件下，视觉伪影以色的形式出现。其中由不匹配的边界条件引起的色漏错误是“隐藏”在高度纹理的区域，这样它将不太明显。因此，作者采取了额外的措施，混合眼睛中存在的无与伦比的边界条件。

在该系统中，作者结合无缝梯度域混合α混合，以避免视觉文物。由于眼睛（皮肤区域）的背景是使用局部颜色转移进行匹配的，阿尔法混合只会在一个适当的α磨砂可以定义为眼睛前景（眼睛区域）时开始工作。否则，从目标的边界或纹理细节（即在图5和图中的闭目）的非平滑转换将导致处理结果失真。该方法，命名为BFGP合成，利用梯度域融合在选择性地结合多种混合效果最好的部分的无缝融合属性，如下：

定义MF的眼睛区域（前景）和MB作为剩余的人脸区域（背景）。首先计算两图像的复合材料，gdbf和gdbb，梯度域融合。对于gdbf，固定前景，使不匹配的边界误差传播到背景；对gdbb，固定背景，使不匹配的边界误差传播到前台的背景。最终的复合将是gdbf和gdbb的α融合，采用作为gdbf背景和gdbb前景（图3）。阿尔法磨砂在这里仅仅平滑掩蔽前景区域MF。在实践中，得到的α-磨砂的引导过滤。

1. 局限与未来工作

该方法有一定的局限性：在某些时候，自动开眼系统不能处理极端差异的姿态和目标和参考之间的照明。输入图像的清晰度越高，该算法产生合理结果的可能性也越低。很明显，系统将面临进一步进步效益和重光照校正。然而，尽管这个限制，对于大多数图像，该系统可以自动恢复和“编辑方法”。在另一方面，一个给定的参考和目标的兼容性全局表达式是一个预测好成绩更好的方法。那是因为该系统不会自动最小化图像对中的表达式差异，但依赖于良好的匹配，而不是从照片收藏。在今后的工作中，这样自动表达式变形可能允许使用更广泛的范围参考图像。

有几种潜在的方法来改善这个系统：

鲁棒性。我们的方法是基于一些组装部件。每个阶段的鲁棒性是整个系统的关键

。例如，姿势扭曲不能处理大的差异姿势；局部颜色迁移可能在drasti产生伪影上不同的光照条件，和混合技术不解决翘曲和重新着色故障以及。每个阶段的鲁棒性，可以改善未来。

人脸图像。操纵面部图像的照明是具有挑战性的任务。该系统，以及其他图像编辑应用程序，可以受益于更好的控制技术人脸光照。

人脸图像的真实感预测。有多种因素影响现实，目前的工作人的感知探索使用手工制作的输入图像的兼容性图像特征。在未来，这将是有益的学习（可能

使用深度学习）。

运行时间。MATLAB实现需要约10s，这是相对缓慢。提高系统的速度对互动和视频应用十分重要

参考文献

[1]  Agarwala. 2007. Efficient gradient-domain compositing using quadtrees.

ACM Trans. Graph. 26, 3, 94.

[2]  Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B.

Curless, D. Salesin, and M. Cohen. 2004. Interactive digital photomon-

tage. ACM Trans. Graph. 23, 3, 294–302.

[3]  Bakhshi, D. A. Shamma, and E. Gilbert. 2014. Faces engage us: Photos

with faces attract more likes and comments on instagram. In Proceedings

of the SIGCHI Conference on Human Factors in Computing Systems.

ACM, 965–974.

 [4] C. Bazin, D. Q. Pham, I. Kweon, and K. J. Yoon. 2009. Automatic closed

eye correction. In 2009 16th IEEE International Conference on Image

Processing (ICIP). IEEE, 2433–2436.

 [5] Bhat, C. L. Zitnick, M. Cohen, and B. Curless. 2010. Gradientshop: A

gradient-domain optimization framework for image and video filtering.

ACM Trans. Graph. 29, 2, 10.

[6]  Blanz and T. Vetter. 1999. A morphable model for the synthesis of 3d

faces.InProceedingsofthe26thAnnualConferenceonComputerGraph-

ics and Interactive Techniquesk. ACM Press/Addison-Wesley Publishing

Co., 187–194.

[7] Bitouk, N. Kumar, S. Dhillon, P. Belhumeur, and S. K. Nayar. 2008. Face

swapping: Automatically replacing faces in photographs. ACM Trans.

Graph. 27, 3, 39.

[8]  Chen, M. Chen, X. Jin, and Q. Zhao. 2011. Face illumination transfer

through edge-preserving filters. In Computer Vision and Pattern Recog-

nition (CVPR), 2011 IEEE Conference on. IEEE, 281–287.