

一种改善光照对深度人脸识别影响的方法

贺 辉,陈思佳,黄 静

(北京师范大学珠海分校 信息技术学院,广东 珠海 519087)

摘要:在人脸识别领域,消除光照变化的不利影响一直以来都是一个难以解决的问题。而与过去的机器学习模型不同,深度学习模型的结构具有和人类视觉神经结构相似的特性。这虽然使模型表现出了非常好的识别效果,但也使模型变得难以解释,以至于以往的人脸光照预处理方法不再可靠。考虑到卷积神经网络具有生物视觉神经的特点,文中在带彩色恢复的多尺度视网膜增强(MSRCR)方法的基础上,结合对比度增强处理,提出了一种类视网膜大脑皮层增强法,以改善基于深度学习的人脸识别模型中光照不均造成的错误识别问题。同时,与基于子空间统计的方法、基于光照不变表示的方法、基于直方图均衡化方法进行了多组对比实验,结果显示该方法比其他方法更有效,可使深度学习模型的识别率显著提高。

关键词:人脸识别;深度学习;光照;视网膜大脑皮层增强

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2019)04-0038-04

doi:10.3969/j.issn.1673-629X.2019.04.008

An Improved Illumination Approach in Deep Face Recognition

HE Hui, CHEN Si-jia, HUANG Jing

(School of Information Technology, Beijing Normal University, Zhuhai, Zhuhai 519087, China)

Abstract: It has always been a difficult problem to eliminate the adverse effects of varying illumination in face recognition. Different from existed machine learning models, the structure of deep learning model is similar to that of human visual nerve. This makes the model show better recognition effect, but also makes it difficult to explain, so that the previous face illumination pretreatment method is no longer reliable. Therefore, considering the convolutional neural network owning characteristics of biological visual nerve, on the basis of multi-scale Retinex with color restoration (MSRCR), combining contrast enhancement processing, we propose a Retinex enhancement method to improve the error identification problem caused by uneven illumination in face recognition model based on deep learning. And compared with the methods based on subspace statistics, illumination invariant representation and histogram equalization, the results show that this method is more effective than other methods, and can significantly improve the recognition rate of the deep learning model.

Key words: face recognition; deep learning; illumination; Retinex reinforcement

0 引言

人脸识别一直以来都是计算机视觉领域的一个研究热点,相比指纹识别、虹膜识别等识别方式,人脸识别有更多优势,因此,基于人脸识别技术的应用也越来越广泛。随着深度学习的兴起,越来越多的领域采用深度学习模型作为主要模型,而在计算机视觉领域,卷积神经网络(convolutional neural network, CNN)成为了最有效的模型之一,人脸识别也不例外,基于卷积神经网络分类模型的方法具有明显优于以往机器学习模型的效果^[1]。神经网络如此强大的一个主要原因是深层神经网络拥有的“万有逼近”能力:深层神经网络可以逼近任意连续函数。而卷积神经网络具有强大的采

样能力,能够自动提取图像集中的主要成分^[2]。然而,虽然人脸识别率已经接近 100%,但是市面上人脸识别设备的应用却很少,主要原因还是模型训练集远远不能覆盖现实中所有的影响因素,而在这些影响因素中,光照是最具代表性的一种。虽然卷积神经网络本身十分强大,在数据集足够好的时候可以几乎不采用任何图像预处理方式,但是当数据集不够全面,或者说缺少足够多的数据时,光照对识别率的影响很大。因此,改善光照对人脸识别的影响对实现人脸识别在工业上的应用有着极其重要的意义^[3-4]。

在光照问题上,近年来并没有提出与 CNN 相结合的方法,主要原因是人为提取高质量的特征十分困难,

收稿日期:2018-04-18

修回日期:2018-08-21

网络出版时间:2018-12-20

基金项目:广东省高校重大科研项目-特色创新项目(自然科学)(2016KTSCX167);广东省自然科学基金(2016A030313384)

作者简介:贺 辉(1979-),女,博士,副教授,研究方向为图像处理与智能分析;陈思佳(1995-),男,硕士研究生,研究方向为深度学习。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20181220.1001.014.html>

并且人为干涉会降低模型提取到的特征的质量^[5],因此现在的主流主张是让模型自主提取特征。例如,特征脸方法^[6]是人脸识别领域内的经典方法,利用 PCA (principal component analysis) 方法计算多张人脸照片的协方差,并求出其特征值和特征向量,接着利用特征值保留最大的若干特征向量,最后利用特征向量对原图像进行投影,这样就达到了保留主成分而降维的目的;基于光照不变表示的方法^[7-8],认为映射到人眼中的图像和光的长波(R)、中波(G)、短波(B)以及物体反射性质有关;局部二值模式法(local binary patterns, LBP)在人脸识别中应用广泛,对光照、年龄、表情等变化都有很强的鲁棒性^[9-10],它通过与周围像素的对比,具有旋转不变性和灰度不变性等特点,但是经它处理后的图像并不符合直觉,换句话说,并不能轻易地由人眼分辨。实际上,经过 LBP 处理后的图像一般不直接用于识别,而是将区域分块直方图连成一个特征向量,放入分类器中做分类,显然这种方法并不适合与卷积神经网络相结合,因为它本身就是一种采样操作,降低了图片的可识别性。

既然 CNN 具有生物视觉神经的特点,那么人为干涉实际上是可以提高模型提取特征的质量的,就像近视眼镜对于近视眼一样。基于这样的考虑,结合直方图均衡化预处理后图像的特点,文中提出了一种类视网膜膜大脑皮层增强法,并通过实验进行验证。

1 算法描述

1.1 类视网膜膜大脑皮层增强法(SRRM)

视网膜-大脑皮层(Retinex)理论^[7]认为世界是无色的,人眼看到的世界是光与物质相互作用的结果,也就是说,映射到人眼中的图像和光的长波(R)、中波(G)、短波(B)以及物体反射性质有关,如式1所示。

$$I(x, y) = R(x, y)L(x, y) \quad (1)$$

其中, I 是人眼中看到的图像; R 是物体的反射分量; L 是环境光照射分量; (x, y) 是二维图像对应的像素位置。

基于 Retinex 理论,有学者提出了 SSR (single scale Retinex) 方法^[8],通过估算 L 来计算 R , 具体来说, L 可以通过高斯模糊和 I 做卷积运算求得,如下:

$$\log R = \log I - \log L \quad (2)$$

$$L = F * I \quad (3)$$

其中, F 是高斯模糊滤波器;“*”表示卷积运算。

通过选择不同的高斯周围空间常数(Gaussian surround space constant)对图像处理有比较大的影响,小的常数对细节和动态区域压缩有比较好的效果,但是整体色彩容易失真,大的常数反之,这也是 SSR 方法的不足之处。

针对这个问题,有学者提出了 MSR (multi-scale Retinex) 方法^[11-13], MSR 使用了多种常数,并用权值的方法将它们混合在一起,如下:

$$\log R = \sum_{i=1}^N w_i \log R_i \quad (4)$$

$$F_i = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{r^2}{\sigma_i^2}\right) \quad (5)$$

其中, σ_i 为高斯周围空间常数; w_i 为每个待混合图像的权值,一般来说:

$$w_i = \frac{1}{N} \quad (6)$$

$$\sum_{i=1}^N w_i = 1 \quad (7)$$

其中, N 为选用高斯周围空间常数的数量。

然而 SSR 和 MSR 对于色彩恢复在灰度上的都有些问题,主要原因在 $\log R$ 恢复到 $[0, 255]$ 色彩空间的方式,也就是恢复到 R 的方式。针对这个问题, Parthasarathy 等提出了带色彩恢复的多尺度视网膜增强算法 (multi-scale Retinex with color restoration, MSRRCR)^[14],如式8~10:

$$\hat{I}_i = I_i / \sum_{i=1}^S I_i \quad (8)$$

$$C_i = \beta \log(\alpha \hat{I}_i) \quad (9)$$

$$R_i = G(C_i \log R - b) \quad (10)$$

其中, C_i 是色彩恢复函数; α 、 β 、 G 都是经验参数; b 是经验偏移量; S 是色彩通道数。

为从根本上消除 MSRRCR 方法导致的图像关键点不明显的缺点,结合直方图均衡化在增强图像对比度上的优点,提出了一种类视网膜膜大脑皮层增强法(similar Retinex reinforcement method, SRRM)。

SRRM 方法同时克服了直方图均衡化方法导致的图像多处变化大的缺点,也即经过 SRRM 处理后的图像具有关键点外的变化度小和有利于目视判读的优点,也即该方法同时保留了灰度增强和视网膜膜大脑皮层法的优点。

算法基本步骤如下:

输入:人脸图像矩阵;输出:增强结果矩阵。

Step1: 将图像转为 RGB 图,并对图像利用 MSRRCR 进行处理;

Step2: 将处理后的图像转为灰度图,进行直方图均衡化处理。

1.2 基于深度学习的人脸识别分类器

基于 CNN 的人脸识别分类器非常多,它们一次次地刷新了 LFW 的记录^[1],甚至有些网络模型拥有非常好的鲁棒性,即便不对数据做过多处理也可以得到非常好的效果^[15]。为验证文中提出的预处理方法的有效性,这里使用一种相对并不复杂的 CNN 结构。

将输入图片作为输入层;第二层是卷积层,卷积核尺寸为 5×5 ,步长为1;第三层是池化层,池化核的尺寸为 2×2 ,步长为2;第四层是卷积层,卷积核尺寸为 5×5 ,步长为1;第五层是池化层,池化核的尺寸为 2×2 ,步长为2;第六层是全连接层,神经元数量为256;最后一层也是全连接层,神经元数量为68,即训练集类别。将最后的输出结果输入到 softmax 函数中做分类。

2 实验及结果分析

文中选用 CMU_PIE 人脸光照数据库作为实验数据集,CMU_PIE 数据集中的 Pose9 是正脸居中裁剪好的人脸数据,一共包含 1 632 张人脸图片,包含 68 个来自多个国家的人的人脸,其中每人有 24 张尺寸为 64×64 的灰度图片,包含 3 张暗光照下不同表情的图片和 21 张不同角度的环绕光照图片。

为了保证数据集倾斜情况的发生,对每张人脸,分别取 19 张图片作为训练集,5 张图片作为测试集,这样训练集有 1 292 张图片,测试集有 340 张图片,训练集和测试集不相交。实验同时对比了文中提出的 SR-RRM 方法与特征脸方法、LBP 方法^[16]、MSRCR 方法、直方图均衡化方法分别对图片进行预处理后的 CNN 的识别效果,对于每张图片,CNN 每次会返回最有可能的预测结果,实验中根据分类器的识别率作为标准。为了保证实验结果的客观真实,对于每种图像处理方法的训练集和测试集,都进行了 10 次随机选取,对于每次选取的数据,又进行了 10 轮神经网络的训练,最终的实验结果是 100 组实验结果取均值,总体技术路线如图 1 所示。各种方法预处理实验结果如图 2 ~ 图 5 所示,最终的人脸识别精度比较如表 1 所示。

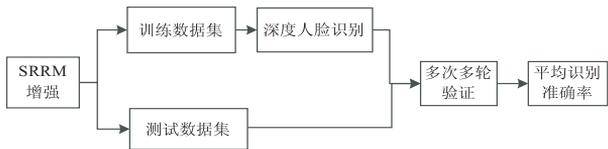


图 1 总体技术路线

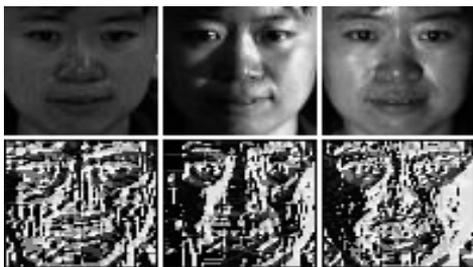


图 2 LBP 处理

(注:第一行是原图,第二行是处理后的图像)

从图 2 可见,虽然 LBP 表现出了强大的人脸识别问题解决能力,但是经它处理后的图像并不符合直觉,换句话说,并不能轻易地由人眼分辨。最后的识别结

果也表明 LBP 不适合与卷积神经网络结合,因为它本身就是一种采样操作,降低了图片的可识别性。从图 3 可见,MSRCR 可以比较好地保留人脸轮廓,消去图中的光照和阴影与皮肤信息,但是对比度不高,一些轮廓细节不明显。由图 4 可见,直方图均衡化预处理也有明显的缺点:变化后的图像灰度级可能会减少,使某些细节不明显甚至消失;均衡化后的灰度范围取决于



图 3 MSRCR 处理

(注:第一行是原图,第二行是处理后的图像)



图 4 直方图均衡化处理

(注:第一行是原图,第二行是处理后的图像)

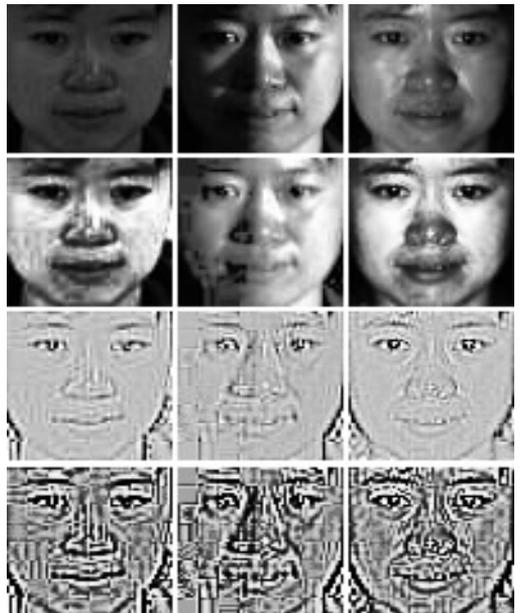


图 5 SRRM 处理

(注:第一行是原图,第二行是经直方图均衡化处理后的图像,第三行是经 MSRCR 处理后的图像,第四行是经 SRRM 处理后的图像)

原图像的灰度范围,因此对灰度范围过小的图像对比度增强的效果有限。而从图5中可以看出,SRRM方法同时保留了灰度增强和视网膜大脑皮层法的优点。

需要特别说明的是,当预处理方法为PCA时,会先将数据集分为训练集和测试集,再让PCA模型对训练数据集拟合,最后再分别对训练集和测试集进行重构预处理,以此来避免预处理方法对测试集的拟合。

表1 多种光照预处理方法与CNN结合后的实验结果

预处理方法	识别率/%
无处理	78.72
PCA,90%主成分,无白化	77.80
PCA,90%主成分,白化	78.53
LBP	56.07
MSRCR	71.58
直方图均衡化	84.13
SRRM	90.74

表1结果显示,SRRM方法相比其他图像预处理方法,在光照处理上拥有更好的效果,明显提升了CNN在光照影响环境下人脸识别的能力。

3 结束语

提出了一种新的光照预处理方法:视网膜大脑皮层增强法(RRM),并与多种典型的光照预处理方法进行了对比实验。实验结果证明,该方法在处理光照不均图像并与CNN结合后的效果远超其他方法,有效地提升了CNN在不均匀光照环境下对人脸识别的能力。更重要的是,提出的CNN和以往的分类器不同,它的识别方式应该符合直觉,也就是说图片应该可以被人眼识别,并通过实验证明了这种想法的正确性,对解释CNN这个复杂的黑盒模型非常有帮助。对于比较极端的光照情况(如半张脸完全被黑暗覆盖),虽然该方法也有复原图像的能力,但是在一些细节上有比较大的瑕疵,针对这个问题,除了一些光照补偿算法外,可以考虑利用人脸的对称性复原人脸,从已经做过的实验结果来看这应该比子空间匹配更有利。

参考文献:

- [1] LEARNED-MILLER E, HUANG G B, ROYCHOWDHURYA, et al. Labeled faces in the wild: a survey [M]//Advances in face detection and facial image analysis. [s. l.]: Springer International Publishing, 2016: 189-248.
- [2] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks [M]//European conference on computer vision. Zurich, Switzerland: Springer International Pub-

- lishing, 2014: 818-833.
- [3] HWANG W, WANG Haitao, KIM H, et al. Face recognition system using multiple face model of hybrid Fourier feature under uncontrolled illumination variation [J]. IEEE Transactions on Image Processing, 2011, 20(4): 1152-1165.
- [4] ZHAO W, CHELLAPPA R, PHILLIPS P J, et al. Face recognition: a literature survey [J]. ACM Computing Surveys, 2003, 35(4): 399-458.
- [5] MAKWANA R M. Illumination invariant face recognition: a survey of passive methods [J]. Procedia Computer Science, 2010, 2(6): 101-110.
- [6] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep learning [M]. [s. l.]: MIT Press, 2016.
- [7] 胡敏, 程天梅, 王晓华. 融合全局和局部特征的人脸识别 [J]. 电子测量与仪器学报, 2013, 27(9): 817-822.
- [8] 王科俊, 邹国锋. 基于子模式的 Gabor 特征融合的单样本人脸识别 [J]. 模式识别与人工智能, 2013, 26(1): 50-56.
- [9] BANIÄ N, LONÄARIÄ S. Smart light random memory sprays retinex: a fast retinex implementation for high-quality brightness adjustment and color correction [J]. J Opt Soc Am A Opt Image Sci Vis, 2015, 32(11): 2136-2147.
- [10] AHONEN T, HADID A, PIETIKÄINEN M. Face recognition with local binary patterns [M]//European conference on computer vision. Berlin: Springer, 2004: 469-481.
- [11] XIE Bin, GUO Fan, CAI Zixing. Improved single image dehazing using dark channel prior and multi-scale retinex [C]//International conference on intelligent system design and engineering application. Changsha, China: IEEE, 2010: 848-851.
- [12] RAHMAN Z U, JOBSON D J, WOODDELL G A. Investigating the relationship between image enhancement and image compression in the context of the multi-scale retinex [J]. Journal of Visual Communication and Image Representation, 2011, 22(3): 237-250.
- [13] MENG Q, BIAN D, GUO M, et al. Improved multi-scale retinex algorithm for medical image enhancement [M]//Information engineering and applications. [s. l.]: [s. n.], 2012: 930-937.
- [14] PARTHASARATHY S, SANKARAN P. An automated multi scale retinex with color restoration for image enhancement [C]//National conference on communications. Kharagpur, India: IEEE, 2012: 1-5.
- [15] GOODFELLOW I J, POUGETABADIE J, MIRZA M, et al. Generative adversarial networks [J]. Advances in Neural Information Processing Systems, 2014, 3: 2672-2680.
- [16] AHONEN T, HADID A, PIETIKÄINEN M. Face description with local binary patterns: application to face recognition [C]//European conference on computer vision. Prague, Czech Republic: [s. n.], 2006: 2037-2041.