



# Contextual Information Based Facial Image Colorization

Yin Junjie

Guangdong Country Garden School, Guangzhou, China

**Email address:**

[oscarfantasywork@outlook.com](mailto:oscarfantasywork@outlook.com)

**To cite this article:**

Yin Junjie. Contextual Information Based Facial Image Colorization. *Asia-Pacific Journal of Mathematics and Statistics*.

Vol. 1, No. 3, 2019, pp. 26-43.

**Received:** September 3, 2019; **Accepted:** November 14, 2019; **Published:** November 27, 2019

**Abstract:** Given no prior knowledge, the process of converting from a grayscale image to a colorful image is an “ill-posed” problem. Most of the previous methods are based on convolutional neural network (CNN), sparse dictionary and user intervention, making colorization either at a huge cost or an arduous work. This paper aims at solving some of the deficiency in previous work, such as methods based on user-intervention require too many human resources, and methods based on machine learning cost too much computational expense. Motivated by this, a novel automatic face image colorization method based on contextual information is proposed by this paper. Our facial image colorization method is based on machine learning. Utilizing the strong correlation between grayscale lightness, texture and color, we first train a joint distribution from our training set, and then solve the color of a targeted grayscale image under multiple constraints including first-order LBP, Second-order LBP, and lightness. Several experiments were performed to show that the proposed method outperforms the previous approaches by offering better authenticity and naturalness. Aiming specifically at facial image colorization, our method manages to achieve convincing results under a relatively small amount of data resources. As a result, this paper achieve desired effect by applying the local binary pattern (LBP) in the field of colorization, and hopefully could be applied in the field of image processing.

**Keywords:** Image Generation, Colorization, Contextual Information, LBP (Local Binary Pattern), Nearby-LBP

---

## 基于上下文信息的灰度人像的彩色化新算法

尹俊杰

广东碧桂园学校, 广州, 中国

**邮箱**

[oscarfantasywork@outlook.com](mailto:oscarfantasywork@outlook.com)

**摘要:** 将灰度图像转化成彩色图像, 它是一个“病态”的过程, 目前, 主要方法是基于字典, 深度卷积神经网络 (DCNN) 等算法和基于人工交互的半自动方式[1-5]。本文针对现有彩色化算法存在的缺陷, 如半自动算法过于需要人力资源, 基于机器学习的自动算法则在计算成本和计算时间上过高等进行了问题的解决。本文在机器学习的基础上, 提出了基于邻域的上下文信息对图像进行训练得到“亮度-灰度-LBP-NLBP”统计人像生成模型; 然后利用灰度值, 1阶, 2阶上下文的约束求解目标灰色图像像素, 对图像进行着色。与已有方法相比, 本文在相当快的时间内使用较少的数据资源且全自动化的实现了一个信服的彩色化的结果, 针对人像, 本文的着色效果更加自然, 具有很好的真实性。综上所述, 本文通过对图像处理的深入研究, 将基于局部二值模式的特征应用在图像颜色重构领域, 取得了良好的效果, 可望在图像处理这块领域获得广泛的应用。

**关键字:** 图像生成, 图片彩色化, 上下文信息, 1阶上下文(Local Binary Pattern), 2阶上下文(Nearby LBP)

---

1. 引言

1.1. 问题的引出

图像的彩色化是指将灰度图片转换成彩色图像，计算机辅助的过程，彩色化主要针对以前的黑白胶卷照片，经典黑白画像，科学展示等已显示出更加视觉性的效果

来增强感官。图片彩色化的主要思想是在一种颜色空间上面传递一个彩色值，将固有的灰度值结合起来形成我们熟知的彩色图片。而彩色化图片相比于灰度图片更具表现力和形象力，在老照片恢复，旧电影上色有着极大的价值。

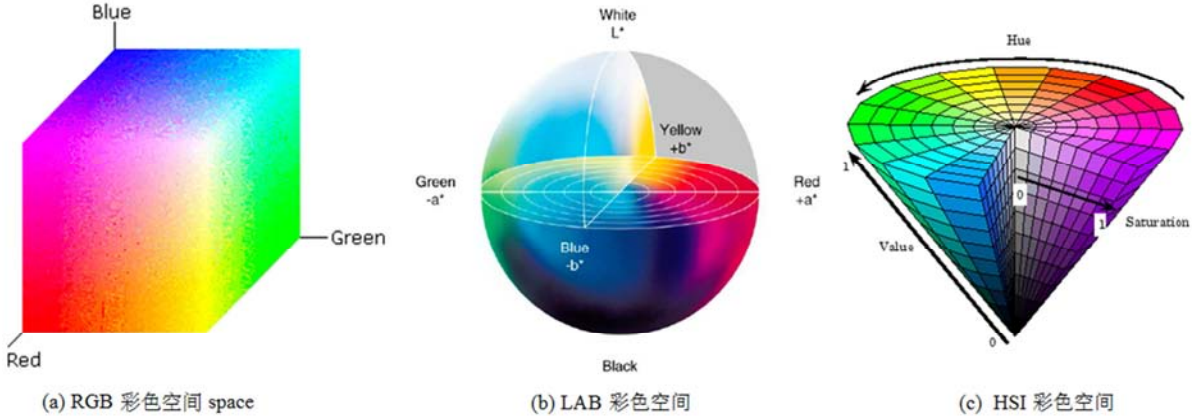


图1 三种颜色空间 Three Color。

1.2. 图像空间的基本原理

本文采取在RGB（Red红，Green绿，Blue蓝）空间进行颜色量化。如图1所示，在图像处理方面的常见图像空间有RGB, HSI和Lab颜色空间，图像彩色化的研究主要是基于这三个颜色空间。不同的颜色空间有不同的特点，如HSI（Hue色彩、Saturation饱和度、Intensity亮度）颜色空间适用于传统的图像处理算法，如卷积、直方图、拟合等。因为RGB在CIE Lab彩色空间之内，所以颜色空间之间可以转换，最后利用不同的颜色空间在用RGB模型里达到不同的效果。Lab和HSI的颜色空间都是有三个分量，一个分量表示非色彩的感知，另外两个分亮表示色彩的感知。在处理图像彩色化的课题当中，本文要彩色化的是一个三维的图像空间，给出的只有一个固定的分量，即HSI里I代表的Intensity（亮度）以及Lab里L代表

的Lightness（光度）。因为在三维空间里在同一个灰度值的情况下，有不同的颜色组合可以组成不同的色彩，所以给出一个分量灰度是没有任何一个唯一的解的，从方程的角度即R,G,B有无数个解，不能满足解的唯一性，因此我们也可以称之为病态的问题（ill-posed problem）。

使灰度图片彩色化一直都是现今的问题，有诸多研究团队已通过Photoshop 的方式来实现彩色化，图2为国外典型的代表Jordan Lloyd的工作[15]，据介绍一张图片完整上彩需要数周的时间，每张黑白图片都要通过五层重重的过程包括将黑白图片重构成上千层的图层，每一个细节都是人工一笔一笔上彩的。而且费用昂贵，多是针对于比较有历史价值的黑白图片，而对于广大群体所可以使用机会极少。

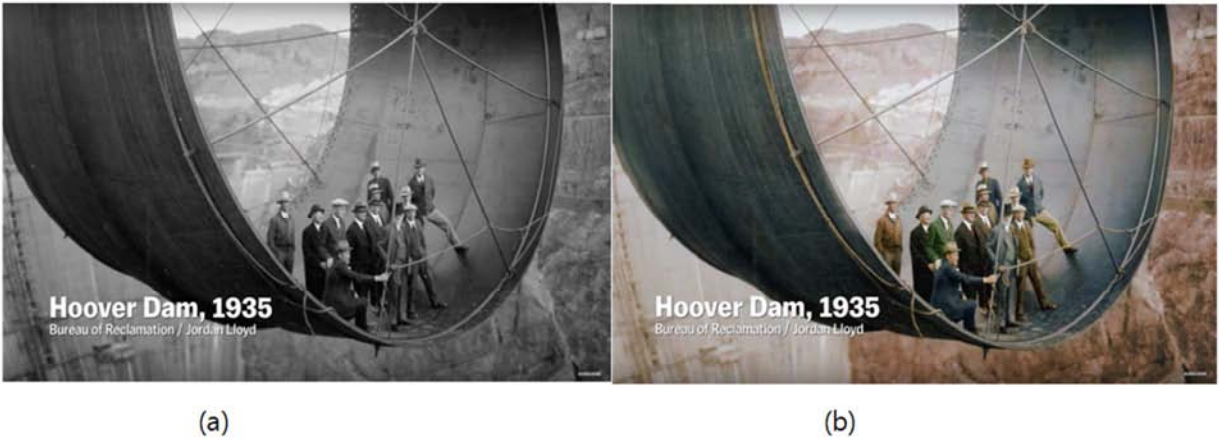


图2 利用Photoshop软件手动上色, 图a为上色前, b为手工上色后。

$$gray(x, y) = 0.2989 * f(x, y, R) + 0.5870 * f(x, y, G) + 0.1140 * f(x, y, B) \quad (2)$$

Definition :  
 $gray(x, y)$  : output image as grayscale image  
 $f(x, y, R)$  : first color channel, red channel pixel value in specific (x, y) coordinates  
 $f(x, y, G)$  : second color channel, green channel pixel value in specific (x, y) coordinates  
 $f(x, y, B)$  : third color channel, blue channel pixel value in specific (x, y) coordinates

图3 彩色图像转换灰度图像关系[8]。

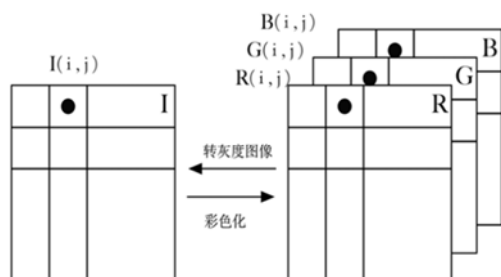


图4 彩色图像与灰度图像的转换关系。

相比过去即使是专业人士使用PS等软件的手动上色也需要消耗大量的时间，运用算法实现自动上色已成为一项新的课题。彩色化不同于图像灰度化，如公式(1)所示，彩色图像转灰度图像的原理是将具有三个分量的彩色图像（如RGB里分别的R,G,B值）通过已有的参数进行矩阵的相乘转成一个灰度值（如Lab的L值）且有固定的运算公式。图3也给出了一个详细的函数表达式。我们可以具体从图4看出给出（R,G,B）三值就可以计算每一个点的灰度值，是一个降维的计算。但若给定I值却是无法求出分别三个矩阵的（R, G, B）三值的。所以彩色化是一个病态的过程。

$$I = (0.2989, 0.5870, 0.1140) \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (1)$$

我们已经知道做一个Lab颜色空间的反函数用一个已知的分量L是无法去推导其他两个分量的，因为有两个未知数所以在彩色空间内可以找出多处与灰度图像完美匹配的点，所以灰度逆运算转换成彩色图像是一个现有图像处理方面的难点。人类视觉系统具有很强的学习性和实验性，所以采用人工着色可以达到一个满意的成果，但明显的劣势使得人工上色面对大量样本的情况下就显得尤为艰难。与之相比，运用机器学习所实现惊人的效果且无需人力资源的消耗的自动化上色模型成为现在彩色化领域的有潜力的研究课题。

### 1.3. 已有工作的综述

相关人员研究的数字图片彩色化算法最主要是在如何得到，处理图像的灰阶和色彩的关系从而区分开来的。能否使原本黑白的照片彩色化成为我们想要的效果，精准的达到使人不能区别开与真实原图的区别，成为了彩

色化算法的关键。目前的彩色化算法可以大致分为两类，半监督以及无监督算法。



图6 左图一为用户添加颜色线条，中间图像为自动扩散的结果[1]。右图为原图。



图5 用户在灰度图上添加彩色线条，算法根据线条进行分割，初步上色，最后通过最优算法得出彩色化图像[3]。

#### 1.3.1. 半自动算法

指给出一副灰度图像，用户可以依据自己的鉴别来人为辅助上色。Levin等提出的优化算法是其中一个典型。通过添加颜色线条，以图5为例，根据用户所添加的颜色笔触设立最优条件[1]，利用算法寻求出最优的图像分割再进行着色。然后算法会将颜色线条的颜色进行扩散，优化，最终得到目标图像。同时也有Yatziv等[2-3]提出基于欧式距离(最短路径)以及颜色混合的算法,图6同样也是用户添加颜色线条，通过计算每个像素到线条的最短距离来决定像素的RGB值为多少。

#### 1.3.2. 全自动算法

指开发一个无用户干涉的算法使得只需要给出一个灰度图像就可以自动的上色。

a) 基于回归和颜色转移的自动化算法：Liu和Deshpande等[6-7]研究主要集中在使用回归得到一个连续且可行的颜色空间，然后通过分类得到颜色值。通过在训练得到特征在与实际的黑白图片寻找相似的特征以达到一个真实的效果。图7



以一篇基于颜色转移的彩色化模型为例[7]，需要提供一张相似的图片作为 引导图通过学习生成一

个回归的柱状统计图，结合灰度图生成最后的输出图。

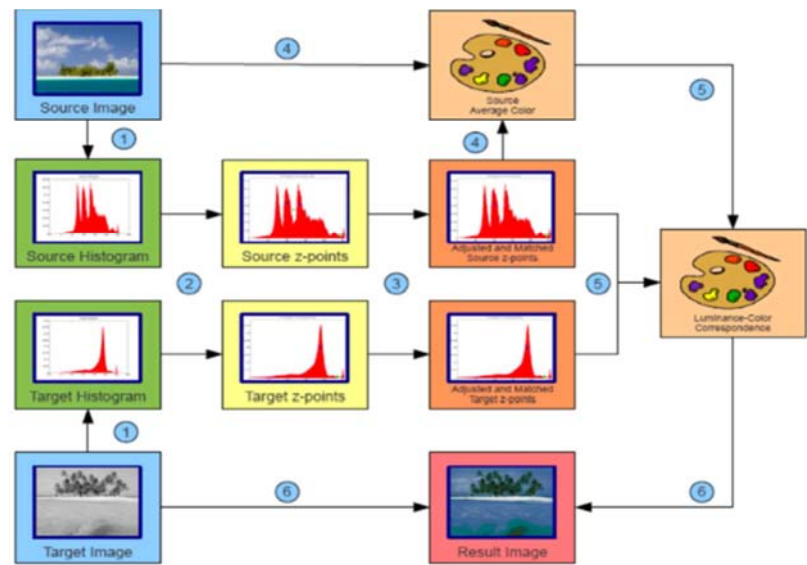


图7 图为利用颜色转移的过程[7]。

b) 基于CNN网络着色：最热门的一类是通过使用深度的卷积神经网络自动彩色化[4-5]，提取特征后都是使用庞大的数据库不断的训练出最后的模型。对于使用深度卷积网络算法，用户在使用上极为方便。给出一个灰阶图像的输入即可，同时随着

使用大量的图像库，得到可行的彩色图片几率会增加。以Zhang[4]的基于深度卷积网络的模型为例，其主要思想是通过大量的学习提取图像特征，在实际着色的过程中通过以前训练得出的相似特征从而上色。

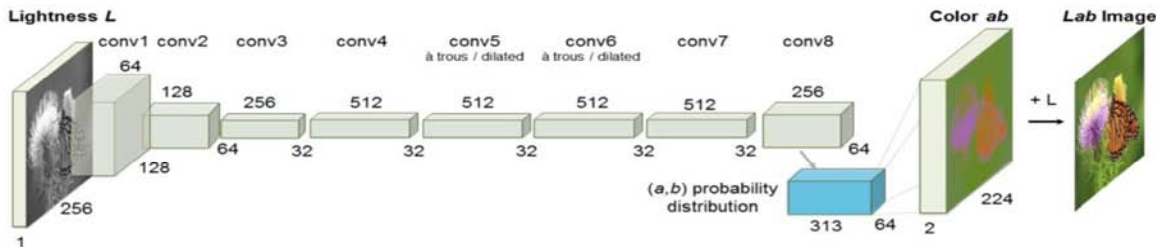


图8 基本卷积网络结构[5]。

c) 基于字典的着色：另一类是则使用稀疏编码使用联合字典[9-10]，利用纹理特征使灰度图像着色。Yan[9]的研究把图片分成一个一个像素格，每一个具有相同特征的像素格子成为一个字典，相同

的亮度和特征即代表颜色也是大致相同的，利用这个假设，故使用大量的训练集同时训练出两个字典，一个基于图片的灰度和特征，另一个基于颜色快最后实现彩色化的效果。

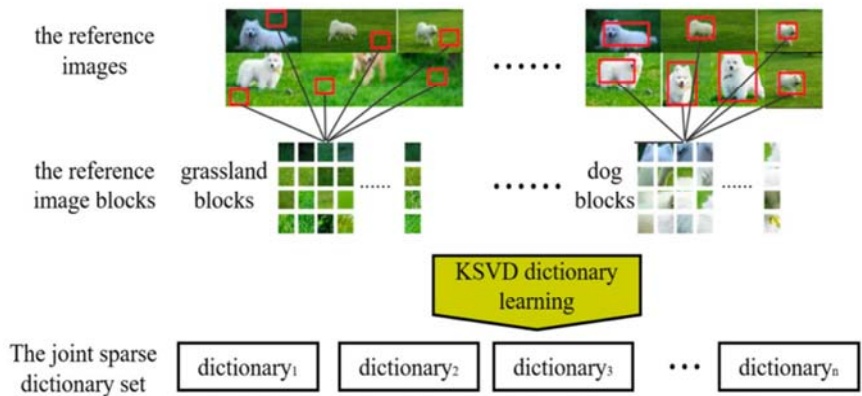


图9 字典学习过程。

#### 1.4. 现有工作的利弊分析

Levin和Lischinski等[1]在基于优化拓展的彩色化研究中,虽然此算法的上色准确率颇高,但是因为需要手动添加线条,不仅对用户所做的笔触要求非常到位,符合图像的特征轮廓,而且耗时相对于其他算法较久,同时不适用于大面积彩色化。而Yatziv[2,3]等对基于距离和色度混合的彩色化研究,属于一个快速算法,可以短时间内完成简单的图像彩色化,但同时因为过多考虑了距离对颜色空间的影响,但忽略了图片可能会出现“颜色聚类”的情况以及其邻域信息。所以在面对特征或纹理比较多颜色较为复杂的图像效果不佳,所上的区域颜色可能会“泄露”到其他区域。

全自动的颜色转移算法[7]虽然不需要与用户直接交互,但因为用户提供了一个相似的图片作为索引,此模型上色速度快且整体的颜色鲜艳。但局限性很强,需要匹配一个特征颜色相近的索引图,且因为对索引图的依赖性强,容易产生上色错误。很多索引图提供的大多是结构上相似却在图像表达的语意上有颇大的区别,导致颜色显得不自然。在图像处理方面非常流行的卷积网络虽然在上色起来极为方便,但所用的图像数据库同时也意味着需要大量的采样,如在[4]里所用的多块顶级GPU需要训练三个周,耗时耗能。其中所用的损失函数也会使得上色后的图片呈现不鲜艳的效果。相比卷积网络,利用字典学习[9,10]的着色算法着色效果更为自然,鲜艳。但同时使用的稀疏编码会导致图像的细节流失以及图像的模糊化,如果使用图像增强技术使得图像恢复原本的特征会使得模型无法正确分类,导致上色错误。

#### 1.5. 本文的动机

鉴于对上述的分析,本文指出半自动算法过于需要人力资源,而普遍基于机器学习的自动算法则在计算成本和计算时间都相当高。为了解决此类问题,本文采取的是一个类似于机器学习的彩色化算法,通过使用适度大小的图像库去训练从而实现一个效果颇佳的彩色图像。与以往使用机器学习或人为辅助上色相比,本模型是基于上下文信息的全自动的颜色重塑模型,侧重于人脸的细节彩色化且使用较小的资源和时耗达到一个逼真

的图像,有助于60-90年代的公众群体简易的使用模型并重新对当时的黑白照片上色,达到足够逼真的效果。

#### 1.6. 文章结构

本篇论文的大致结构如下,第二部分主要介绍使本文提出基于0阶,1阶,2阶上下文信息采取最大概率从而着色的彩色化算法。在第三部分,对本文彩色化算法进行了实验从而证明本文开发的彩色化算法的效果以及其实用性。在第四部分则对全文所设计的进行了总结以及评估。

### 2. 基于上下文信息的颜色重塑工作

#### 2.1. 本文工作

上下文信息通常被理解为:能应用能够影响图像中的对象的某些信息。在机器视觉里面上下文信息的定义则是通过捕捉图像里面某一个区域的领域内的数据而得到的信息称之为上下文信息。而本文在随后提出的LBP等纹理特征模式也属于上下文信息的一种。

本文的主要研究目的是基于上下文信息特征提取的图像处理应用,核心思想是采用LBP或类LBP(NLBP)对图像库里的图像的局部信息进行组织与统计,最后通过随机或最大概率求出每一个像素值的RGB三通道的颜色值从而上色。由于现今普遍都采用机器学习的方法,图像处理过程需要经过特征提取、模型学习以及得出结果这三个步骤,一般情况下都需要大量的采集样本进行训练,导致了算法的时间复杂度非常高,许多研究彩色化有成效的自动算法难以做到实时,而且还可能无法广泛应用。基于对传统半自动算法需要人为交互使得广大群体难以使用且耗费用户的精力以及时间,自动算法需要花费大量的时间和数据处理工作的考虑,本文的这项工作将着重解决于这些问题,使用较少的数据资源全自动化的达到一个信服的彩色化的效果。本项工作提出了一个新颖的重塑颜色的架构并细分出了两种分别基于局部二值信息(LBP)和次局部二值信息的算法。通过同时添加两至三个约束——灰度值以及两种上下文信息找出在彩色空间内与其匹配的点,自动化的将一张灰度图转化成彩色图。

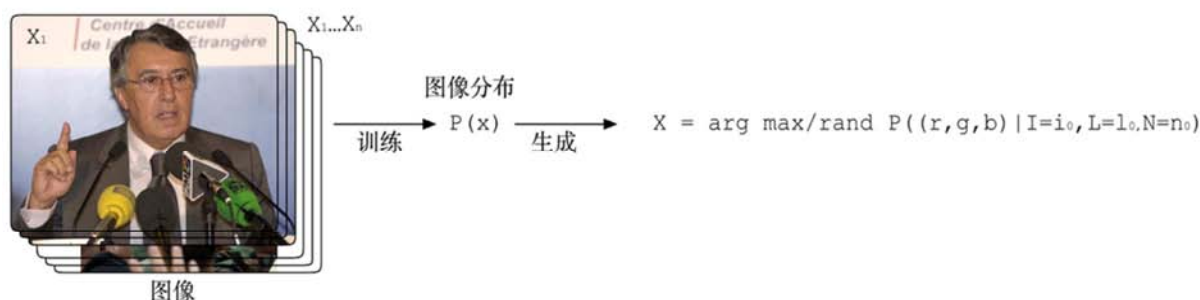


图10 本文图像生成的原理。



图11 图为随机生成的噪声。

## 2.2. 图像生成

图像是信息传递的主体，图像生成则是图像处理领域的最基本的研究课题之一。图像生成一般的彩色图像信息具有三个分量以表示其颜色值以及亮度值，而灰阶图像只存在亮度值故只有一个分量。公式(1)已阐明从彩色图像到灰空间图像是有一个固定的运算公式，但因为存在两个未知的分量，是无法求解的，对于一副 $m \times n$ 的灰度图像它的彩色空间为 $256^{(m \times n)}$ 。而彩色图像因为有三个通道所以空间为 $256^{(m \times n \times 3)}$ ，非常巨大。通常的随机生成模式在这个空间中生成的图像大多数是没有实际意义的，如图11所示。

为了使上彩结果获得意义且有一个有效果的图像，必须添加某种或多种特殊的条件来保证转化的过程中是只有一定的可行的解。本文采取的特殊条件的方式为添加约束，在添加约束后使得解的范围收敛到一定的小的范围，最后在取其中满足解的彩色值。根据这种模式要生成一幅有意义的图像需要一定的数据进行训练。根据图10的生成模式，本文分别设计了三个基于上下文信息的人像彩色化算法。

本研究的原理如公式(2)所示， $X$ 为目标彩色图像的每一粒像素，求满足于 $L=l_0, C=C_0, N=n_0$ 的变量达到最大值或随机的取值。最后通过每一个像素的自动化上彩，RGB三通道分别赋予颜色，最后生成图像。其中 $\max$ 代表取分布中最大概率， $\text{rand}$ 则指随机取符合分布中的值。

$$\chi = \arg \max_{\text{rand}} P((r, g, b) | L = l_0, C = c_0, N = n_0) \quad (2)$$

s.t.  $l_0, (r, g, b)$ 满足公式(1)

其中 $l_0, c_0, n_0$ 分别代表单个像素中的灰度值，LBP值，和NLBP值。

## 2.3. 基于0阶LBP的灰度图像着色算法

图像颜色重塑属于图像处理这一大块领域，彩色图像一般采用RGB空间，其像素值可以用矩阵量 $C = [r, g, b]^T$ 表示。简单而言，图像着色是用残缺的信息恢复出完整的信息的过程。通过对大量彩色人脸的图像分析和总结，本文发现对一张彩色图像本身是有部分规律可循的。一下分别是两种可靠的假设：

1. “若灰度空间邻域的灰度值相似，则彩色值也会大致相似”。在自然人脸的灰度图像中，肤色的过渡较为平滑，色调的分布也相对均匀，因此若图像像素的亮度值相近，则颜色一般会比较相似；
2. “局部的纹理信息具有相似的颜色”。在自然人脸图像中一类纹理通常对应相似的颜色，如人的嘴唇多为红色，眉毛多为黑色，而眼睛则多为眼白和黑色眼珠。

本文首先只添加一种约束条件，即灰度图像中的灰度值与颜色信息值的相关性，本文首先遵循第一个假设，将由亮度的参数结合导颜色信息之中，得到一个二维的分布，在从此分布随机映射一个满足分布的颜色值。从而恢复除目标图像的颜色信息。具体流程如图12所示。

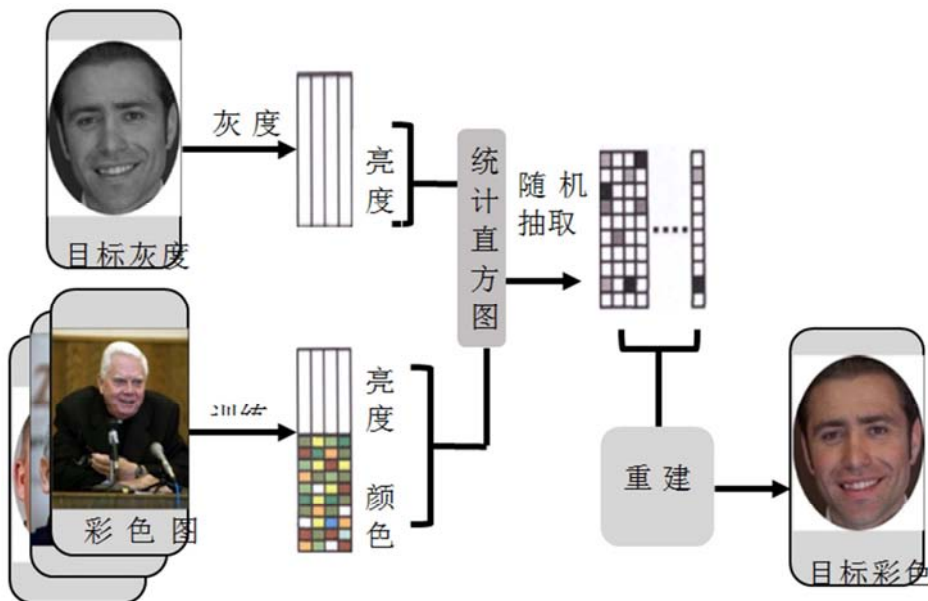


图12 利用0阶LBP的自然图像着色流程图。

具体实现可以分为以下几个步骤：首先依据彩色图像库训练得到每一个灰度值所对应的 RGB 颜色分布，本文采用的图像库一共有 5k+，每张图片是包含人脸的。同时采用本文采用二维的统计直方图的方式进行实现。

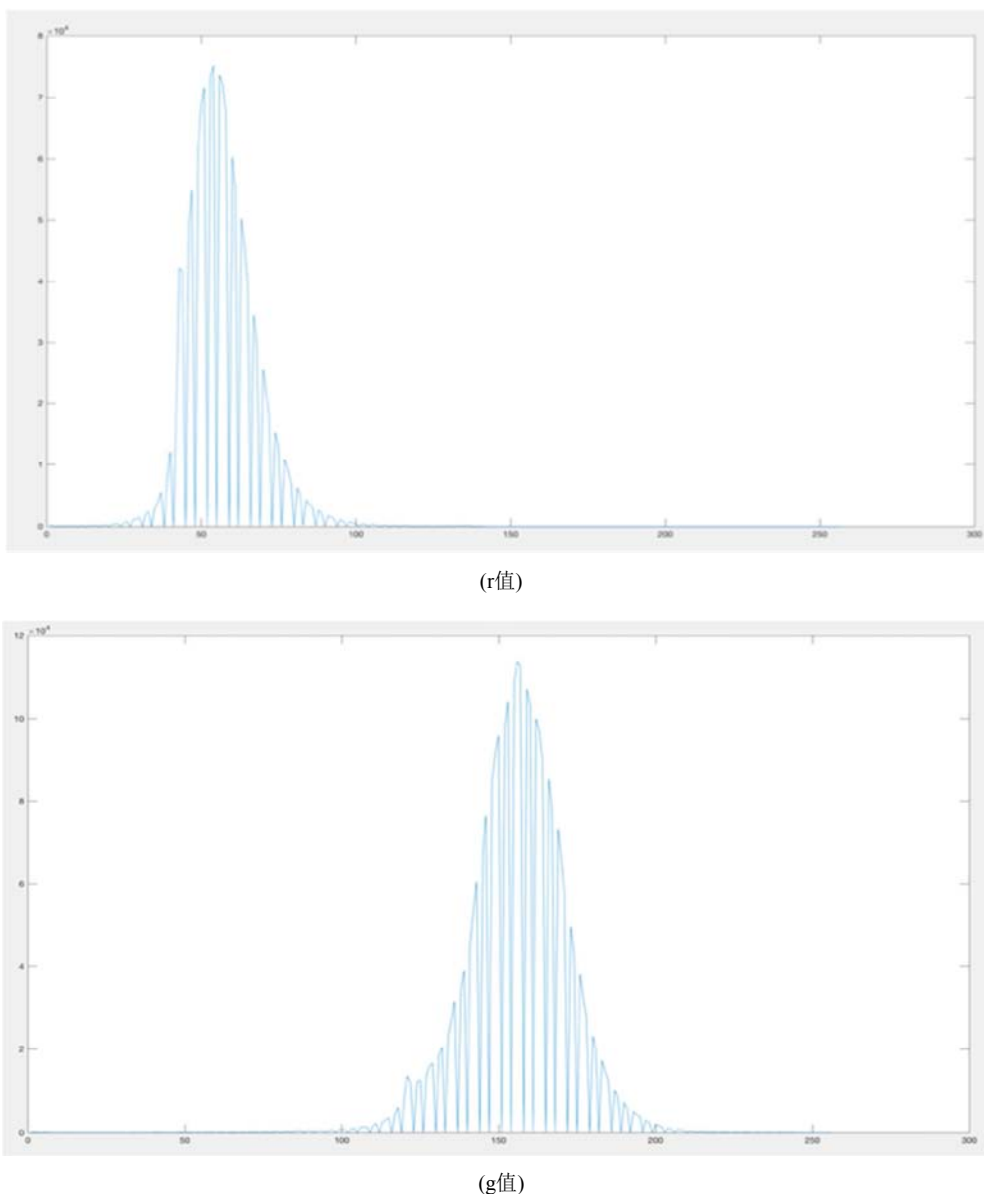


图13 统计灰度值与颜色值直方图的r,g彩色通道截面（通过固定其他两个颜色值所得一维截面）。

于是利用公式(1)使每张图像库中灰度值所得到的统计直方图，在与对应的灰度值里根据训练所得经验取满足分布的RGB颜色值，通过概率随机抽取此RGB值，求解目标灰度图像，从而对图像进行着色。公式(3)展示了本文的求解原理。

$$\arg \max_{r,g,b} P(r(x,y)g(x,y)b(x,y) | L = l_0) \quad (3)$$

算法 1. “灰度”的0阶上下文约束算法（训练/生成）  
训练阶段Training Phase

1. Input: 读入K个m\*n大小的图像P;
2. lr: 256\*256 的向量矩阵;

3. lg: a 256\*256的向量矩阵;
4. 初始化 lr, lg;
5. For i = 1 : file\_num do //file\_num 为图像总数
6. I = RGB2GRAY(P); //图像P转成灰度图像I
7. For j = 1,...,m do
8. For k = 1,...,n do //m\*n里的每一个像素格内
9. lr(i, r) = lr(i, r) + 1; //更新矩阵内第j行k列的元素
10. lg(i, g) = lg(i, g) + 1;
11. End for
12. End for
13. End for
14. 将训练所得分布D1, D2存档



15.  $D1 \in \text{RGB}$  通道里r的分布

16.  $D2 \in \text{RGB}$  通道里g的分布

训练阶段Training Phase

1. Input: 读入K个 $m \times n$ 大小的图像P;
2. lr:  $256 \times 256$  的向量矩阵;
3. lg: a  $256 \times 256$ 的向量矩阵;
4. 初始化 lr, lg;
5. For  $i = 1 : \text{file\_num}$  do //file\_num 为图像总数
6.  $I = \text{RGB2GRAY}(P)$ ; //图像P转成灰度图像I
7. For  $j = 1, \dots, m$  do
8. For  $k = 1, \dots, n$  do //  $m \times n$ 里的每一个像素格内
9.  $\text{lr}(i, r) = \text{lr}(i, r) + 1$ ; //更新矩阵内第j行k列的元素
10.  $\text{lg}(i, g) = \text{lg}(i, g) + 1$ ;
11. End for
12. End for
13. End for

14. 将训练所得分布D1, D2存档

15.  $D1 \in \text{RGB}$  通道里r的分布

16.  $D2 \in \text{RGB}$  通道里g的分布

生成阶段Generating Phase

1. Input: 读入 $m \times n$ 大小的灰度图像I
2. 初始化 imn为 $[m, n, k]$ 的三维, RGB格式的矩阵;
3. For  $i = 1, \dots, m$  do
4. For  $j = 1, \dots, n$  do //  $m \times n$ 里的每一个像素格内

5. 利用最大/随机概率求灰度图像I $[i, j]$ 里的灰度值z所满足的分布D1的r值;
6.  $r\text{值} = \max/\text{rand}(D1(z,:))$ ;
7. 赋r值在imn的 $[i, j, 1]$ 元素中;
8. 利用最大/随机概率/求灰度图像I $[i, j]$ 里的灰度值z所满足的分布D1的g值;
9.  $g\text{值} = \max/\text{rand}(D2(z,:))$ ;
10. 赋g值在imn的 $[i, j, 2]$ 元素中;
11.  $b\text{值} = (\text{gray1}(i, j) - 0.3 \cdot r - 0.59 \cdot g) / 0.11$  通过公式计算求的b值
12. 赋b值在imn的 $[i, j, 3]$ 元素中;
13. End for
14. End for
15. 返回 imn.

彩色图像的初始重建结果有分别几个不足：一是整体的颜色饱和度不够高，因为图像库有部分灰度图像可能学习的效果降低。二是整体的色彩平滑度不够自然，因为在颜色重塑的过程中只有一个约束在限制此病态问题的求解，所以随机的采样只能保证全局的颜色贴近肤色。三是局部特征的颜色没有匹配，没有添加提取特征的约束。实验结果表明，使用单一的约束的方法求解彩色的另外两个分量是不够的，但可以证明达到的效果是一个整体平滑的目标彩色图。根据这三个问题，本文跟进的提出了基于二值特征模式的自动彩色化算法，用局部特征作为约束，在寻找相对分布的同时考虑到邻域的信息进行颜色重建，增加了局部图像块的颜色饱和度，并且提高了邻域像素之间的平滑度。

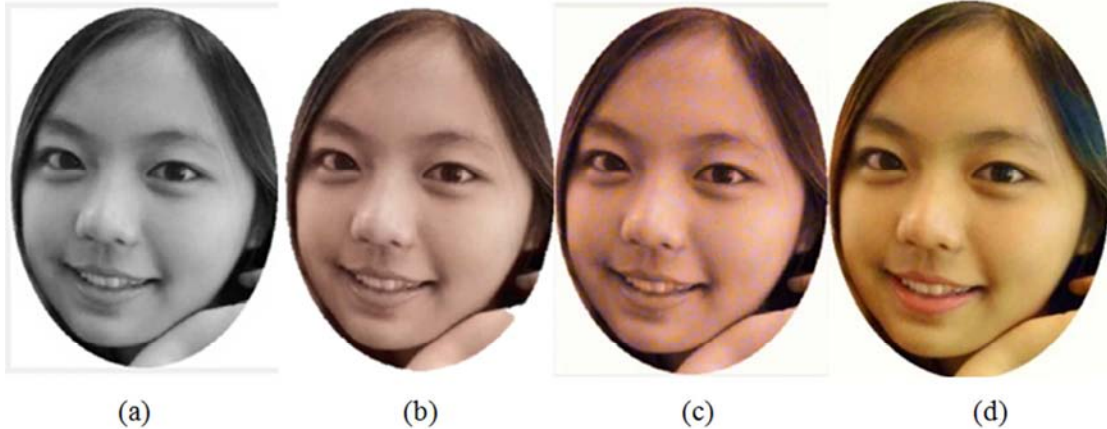


图14 (a)为灰度图像 (b)为Zhang等[4]所采用的卷积网络的结果 (c)为0阶LBP的结果 (d)为彩色原图。

## 2.4. 基于LBP与灰度值的一阶LBP灰度图像自动着色算法

### 2.4.1. 局部二值模式LBP

LBP指局部二进制模式 (Local Binary Pattern)，它是首先由T. Ojala, M. Pietikäinen, 和D. Harwood [14]在1994年提出的纹理特征提取的一种算子。它具有旋转不变性以及多分辨率等显著优点，并且所提取的特征是图

像的局部的纹理特征。相比于灰度值不同的是，局部纹理的特征可以更加有效的体现其提取的特征值和颜色值的关系。正如前文假设2所提出的，局部纹理的信息会具有相似的颜色。因为计算不仅简单，而且效果较好，因此LBP在许多图像视觉处理方面有着广泛的应用。但是还没有投入在彩色化图片这个领域之中，本文将会采用这种特征提取方式以此为第二个约束，从而进行上色。



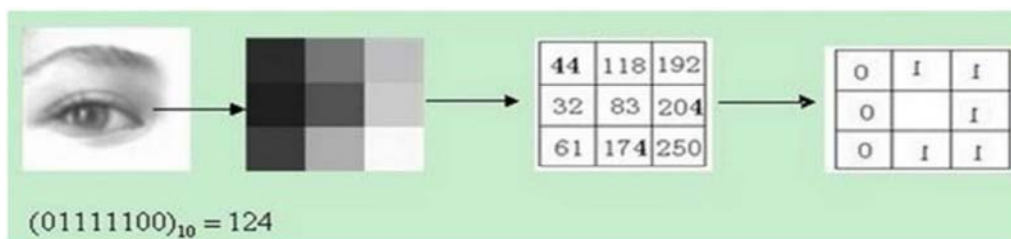


图15 图像转换成LBP矩阵的过程。

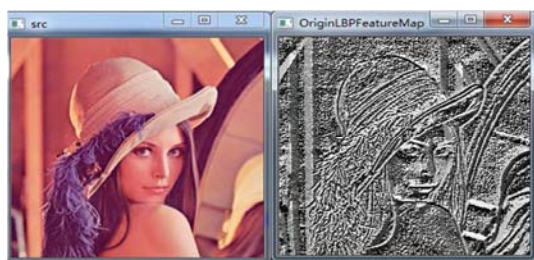


图16 左图为原图，右图为计算LBP后生成的LBP图像。

LBP的计算方法是在一个定义的3\*3的邻域内，在邻域内的中心点作为阈值，阈值与左上角的值以此顺时针方向依次进行比较，若相邻像素点大于中心像素点的值，则标记为1，否则标记为0。这样就使得在3\*3的邻域内比较八次，再把这八个二进制数按顺序按原先的顺序排好，就生成了一个八位的二进制数字。而这个数字就是中心像素的LBP值，LBP有2<sup>8</sup>种可能性，因此也就有256种可能性。中心的LBP值也就反映了其邻域的纹理特征。需要注意的是如果是彩色图的必须要先转成灰度图在进行LBP的计算。LBP的数学描述如下：

$$LBP(x_c, z_c) = \sum_{p=0}^{p-1} 2^p s(i_p - i_c) \quad (4)$$

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{else} \end{cases}$$

#### 2.4.2. 基于1阶上下文的三维颜色重构

对于前章的问题，本文必须研发一个更完善的重塑结构，既能体现出彩色图像的特点信息，还可以在灰度值与彩色值中建立某种关系。我们可以很容易地想到亮度信息是两种不同模式的图片共同具有的，但是这种训练出的分布在实际应用表现不是特别理想，会导致饱和度不高的着色结果。

在自然人脸图像中，纹理信息可以将一张图像显著的分成不同区域的颜色，但是仅仅使用“纹理-颜色”的训练分布可能会导致重新上色的颜色内容紊乱，因为纹理信息总归还是一种抽象的特征，通过LBP就找出颜色与灰度值的关系是不现实的，如果利用这样的分布反过来对眼色信息进行重建会有损图像的原始内容。

综上所述，为了解决先前所描述的问题，本章节提出了基于LBP与灰度值的一阶LBP彩色算法，是一个“亮度-上下文信息-颜色”的三维分布。三个特征向量分别代表由255个灰度值的向量，255个LBP值的特征向量以及每个像素的RGB三通道的颜色值排列构成，由亮度和纹理共同充当传递颜色时候求解的约束，如图17所示。

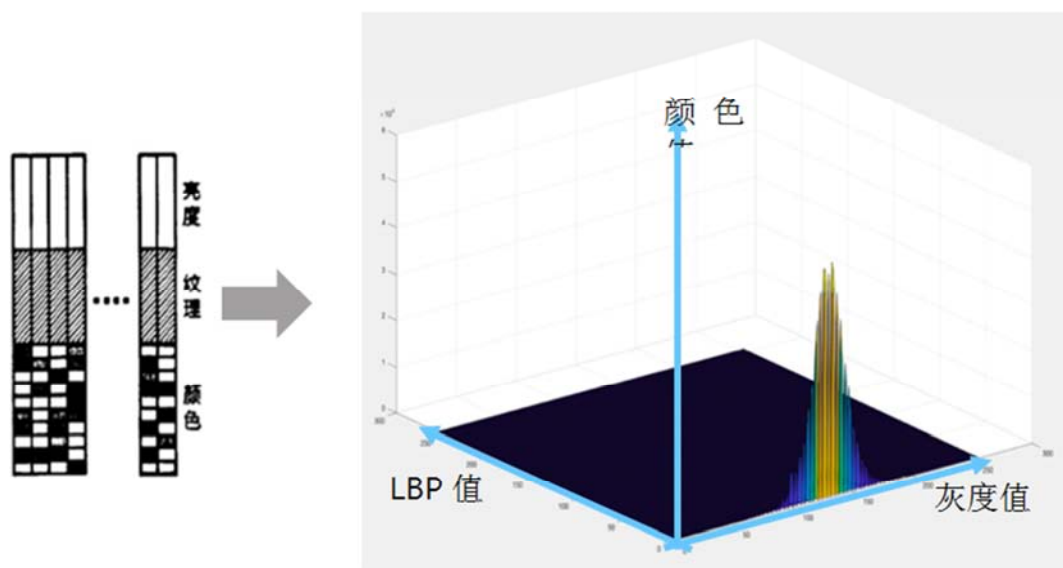


图17 亮度-局部二值特征(LBP)-颜色的三维分布。

明确了本章的基本思想过后，还有几点需要注意：

首先，虽然在图像着色领域如Zhang[4]和Levin[1]的研究都是在LAB的颜色空间内实现的，但本章采取的是

用LBP和灰度值找出可行的RGB像素值的分布，RGB空间具有很好的线性叠加性，即意味着RGB可以很好的不变性，在不同空间所展示出来的都是一样的内容。

其次，与上次不同的是本章采取了一个与上一章节所不同的人脸库，其中一共只有5k+张全彩的人脸图，像素均为256\*256大小。

在确定基于LBP的条件概率分布之后，理论上会有这样的公式模型，在任意R,G,B值表示符合灰度值 $g = g_0, L = L_0$ 的条件概率为下：

$$P(r, g, b) | L = l_0, C = c_0) \quad (5)$$

其中 $l_0, c_0$ 分别代表像素的特定灰度值与LBP值。根据上式表达，R, G, B值会分别满足于分布，最后再将数个对应值（满足条件概率的值）的其中一个抽取出来做RGB的颜色值。

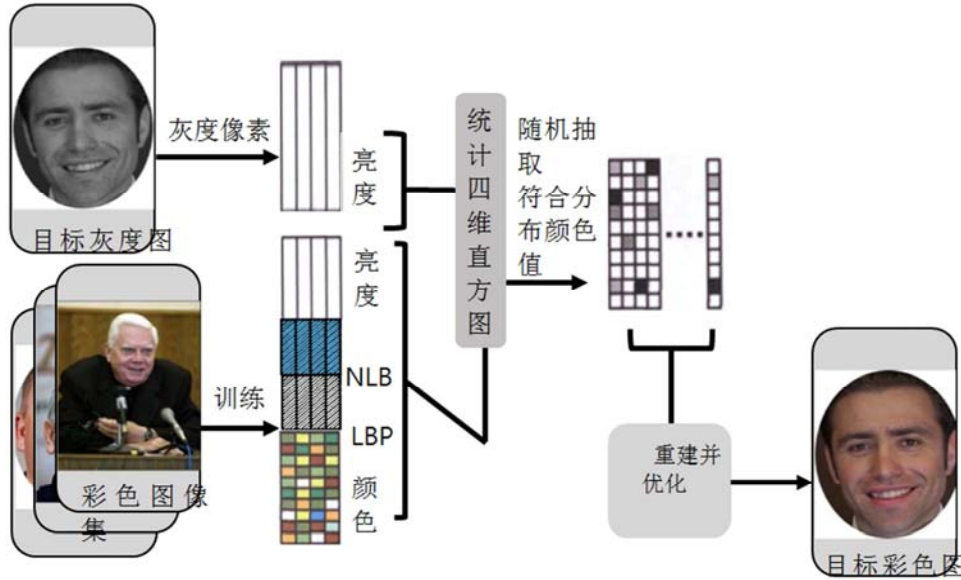


图18 基于二阶上下文信息人脸图像着色流程图。

着色的流程图如图18所示，由于二阶上下文信息彩色化算法流程大致相同，唯一不同只是在训练过程中了NLBP的特征因子，所以将此流程图作为统一流程图。

该方法具体实现如下：

1. 对一副灰度图像进行自动着色，要给出一副灰度图像；
2. 把含有亮度信息的图像读入，进行判断每一个像素块所对应的LBP值与灰度值。在这个像素点上的亮度值判断是否满足于所训练出来“灰度-局部二值特征-颜色”的三维分布。利用公式(5)进行计算RGB值得颜色信息，即 $f_1(g, l) = R, f_2(g, l) = G, f_3(g, l) = B$ ；
3. 进行图像重构，分别RGB三通道每一个通道对应一个 $[m \times n]$ 的向量矩阵，将所计算出的RGB三通道的三个值分别传递到每个矩阵的元素中。
4. 最后将R,G,B三个通道的矩阵值进行合成，就可以得到一副重塑的目标彩色图像。

该方法的试验结果如下图19所示，由实验结果可得知该方法取得不错的彩色化的效果。如图所示，该方法彩色化效果明显要优于利用Lizuka[5]彩色化的效果。

算法 2-3. “灰度-LBP-NLP”的二阶上下文算法（训练/生成）

训练阶段Training Phase

1. Input: 读入K个 $m \times n$ 大小的图像P;
2. Cr:  $64 \times 64 \times 64 \times 32$  的四维向量矩阵; //代表着RGB通道的r值分布

3. Cg:  $64 \times 64 \times 64 \times 32$  的四维向量矩阵; //代表着RGB通道的g值分布
4. 初始化 Cr, Cg;
5. For a = 1 : file\_num do //file\_num 为图像总数
6. I = RGB2GRAY(C); //图像C转成灰度图像I
7. For j = 1, ..., m do
8. For k = 1, ..., n do //m\*n里的每一个像素格内
9. lbp = gray2lbp(I)
10. nlbp = gray2nlbp(I) //将灰度图像I转为LBP图像
11.  $Cr(i, r, lbp(i, j), nlbp(i, j)) = Cr(i, r, lbp(i, j), nlbp(i, j)) + 1$ ;
12. //i为灰度图像I j行, k列的元素
13.  $Cg(i, r, lbp(i, j), nlbp(i, j)) = Cg(i, r, lbp(i, j), nlbp(i, j)) + 1$ ;
14. //r, g 分别为C的两个RGB通道
15. //更新矩阵内第j行k列的元素
16. End for
17. End for
18. End for
19. 将训练所得分布D5, 64存档
20. D5 ∈ RGB 通道里r的分布
21. D6 ∈ RGB 通道里g的分布

生成阶段Generating Phase

1. Input: 读入 $m \times n$ 大小的灰度图像I
2. 初始化 imn为 $[m, n, k]$ 的三维, RGB格式的矩阵, 新图像;
3. For i = 1, ..., m do
4. For j = 1, ..., n do //m\*n里的每一个像素格内

5. 利用最大概率/随机概率求灰度图像 $I[i, j]$ 里的灰度值 $z$ , LBP图像里的LBP值 $l$ , NLBP图像里的 $n$ 所满足的分布 $D5$ 的 $r$ 值;
6.  $r$ 值 =  $\max/\text{rand}(D1(z, l, n : ));$
7. 赋 $r$ 值在 $\text{imn}$ 的 $[i, j, 1]$ 元素中;
8. 利用最大概率/随机概率求灰度图像 $I[i, j]$ 里的灰度值 $z$ , LBP图像里的LBP值 $l$ , NLBP图像里的 $n$ 所满足的分布 $D6$ 的 $g$ 值;
9.  $g$ 值 =  $\max/\text{rand}(D2(z, :));$
10. 赋 $g$ 值在 $\text{imn}$ 的 $[i, j, 2]$ 元素中;
11.  $b$ 值 =  $(\text{gray1}(i, j) - 0.3 * r - 0.59 * g) / 0.11$  通过公式计算求的 $b$ 值
12. 赋 $b$ 值在 $\text{imn}$ 的 $[i, j, 3]$ 元素中;
13. End for
14. End for
15. 返回  $\text{imn}$ .

我们可以得知, 本文所训练出来着色图片的可以把整个人的特征体现出来, 即肤色是符合人的基本判断的。考虑到现有的灰度图像着色, 本文选取了自动化算法中最具有代表性的方法, 使用卷积网络的方法与本文的研究进行比较。可以看到在Lizuka等的方法中, 所得出来的彩色结果会有一定程度的互相干扰, 导致上色出现了很大部分的上色错误, 单从肉眼来看, 可以明显看出色差是与原图有很大差别的, 除了脸颊其他部分并不是很理想。这是因为基于卷积神经网络的着色方法在针对一些没有训练过的图像具有劣势, 尤其是在当整体灰度的处在一个相对一致的情况下, 卷积的方法会有多种不同的上色可能。在这一点上, 给予上下文信息的结果有明显的优势, 有一定纹理特征的基础进行着色, 所以对像素的颜色着色正确率会远高于Lizuka等方法。

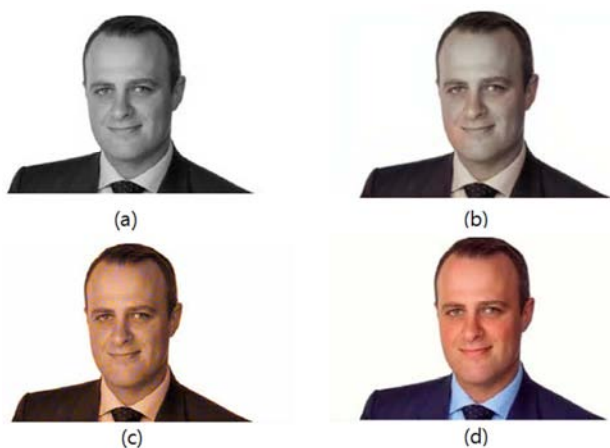


图19 a为灰度图像 b为Yatziv等[5]所采用的卷积网络的结果 c为1阶上下文信息的结果 d为彩色原图。

如图20所示, 本文从先前使用单一约束灰度值的结果与使用双重约束的结果进行了对比, 可以看到重建结果中无论是从纹理还是颜色来说, 使用双重约束所得的分布远比单一约束的效果要好。可以看到第二张图的邻域颜色都属于肤色的范畴, 最终获得了一个饱和度较高的着色结果。通过添加LBP约束, 最初大片的呈灰绿色

的肤色在添加约束后颜色转换成肉色, 饱和度增高。整幅图的平均色差明显的有降低。



图20 两种方法比较图a 为使用0阶约束所得到的结果 b为利用二阶约束所得到的结果。

## 2.5. 基于二阶上下文信息的自然灰度图像着色算法

### 2.5.1. 次局部二值特征模式(次LBP)

在本文的前段已经讲到LBP值, 指的是在一个 $3 \times 3$ 的邻域内, 在邻域内的中心点作为阈值, 阈值与左上角的值以此顺时针方向依次进行比较, 最后形成一个8位, 256种可能的LBP值。

$$\begin{bmatrix} 1 & 2 & 3 \\ 8 & LBP & 4 \\ 7 & 6 & 5 \end{bmatrix} \quad (6)$$

$3 \times 3$ 的LBP值相应的是考虑到的是其像素块每一个相邻的像素块, 特征值也就可以体现出来。但从一张 $256 \times 256$ 的图像角度来说, 使用单一LBP值所计算出来的邻域信息仅能表现出一部分的特征信息, 可以说使用LBP所计算出的上下文信息是“不完全的”。本研究设计了一个“邻域”的局部二值特征模式, Nearby Local Binary Pattern(NLBP)。如下矩阵所示, 指在一个 $5 \times 5$ 的邻域内, 与LBP原理类似, 以邻域内中心为阈值, 阈值与邻域与左上角的值进行比较, 然后在与其对应的正上方进行比较, 以此为序依次比较。这样在 $5 \times 5$ 的邻域内比较八次, 就可以产生另外一组八位的二进制数, 也是256种可能。

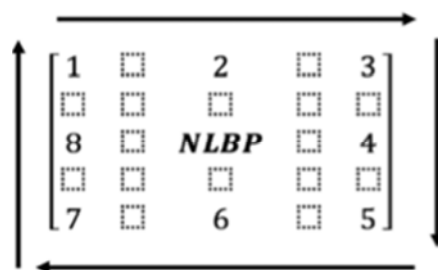


图21 NLBP局部特征概览。

不同于标准的LBP, NLBP描述的是一个范围更加广的局部特征。这种特征基于图像的上下文信息, 考虑到的是一个范围更加广的邻域, 以求一个更加完善的上色效果。添加一个更广的邻域的约束意味着所训练出分布会在LBP, 灰度值中间多一条纽带。使得上色效果更平滑, 上色效果也更为贴近真实。



### 2.5.2. 基于二阶上下文的四维颜色重构

前章所提出的基于一阶LBP的图像在自然人像着色块已有显著的进步, 添加了邻域信息的约束后, 有明显的改善效果。但是3\*3的邻域信息在偌大的图像中, 单单只考虑一种小规模邻域信息会显得有些乏力, 结果会导致图像的局部特征部分上色可能会出现“噪音”的情况, 即在随机取值时会有概率使单个像素粒上色错误。针对这个问题, 为了使得图像更佳富有形象力, 本文研发了一个覆盖面积更广的局部二值特征: 次局部二值模式 Nearby Local Binary Pattern(NLBP)。目的是使的纹理不仅考虑到周围的8个值, 同时还考虑到周边的邻近值。



图22 LBP与NLBP因子。

如图22所示, 蓝色以N开头的符号代表着NLBP所考虑的纹理信息, 而红色以L开头则为LBP所考虑的纹理信息。相比于之前的算法, 若添加一个次LBP值, 产生的是一个二阶的纹理特征, 同时考虑到相邻, 周边的特征。

综上所述, 本章将结合本文所研究的算法与次局部二值模式, 提出“灰度-LBP-次LBP-颜色值”的四维分布。四个特征向量 分别代表由255个灰度值的向量, 255个LBP值的特征向量, 255个NLBP值的特征向量以及每个像素的RGB三通道的颜色值排列构成。这是一个四维的分布, 由亮度和二阶LBP纹理共同充当传递颜色时候求解的约束, 如图18（第十页）所示

本章的主要思想是通过增加约束使的求解过程中的解集具有收敛性, 使的可行的颜色增多, 错误的颜色减少, 最终达到一个平滑度高, 逼真度高的自然着色图像。本章结合过往实验经验也做出了以下几点调整:

其一, 创造一个四维的分布是需要数以百万计的彩色图像才能训练出来一个相对稀疏的分布, 256\*256\*256\*256的分布意味着包含 $2^{32}$ 个位置的若干立方体, 考虑到本文使用的训练库的大小, 同时考虑到训练的时耗以及效率, 本文将四维的分布缩小, 将每四个邻近值划分为单一的值, 因为NLBP所计算的周边信息也较为广泛, 所得到的值也比较统一, 所以将每8个邻近的NLBP值划分为一个统一的价值, 以减少训练时间, 同时得到高度相似的结果。

二, 在确定基于二阶LBP的四维条件概率分布之后, 数理上会有这样的公式模型, 在任意R,G,B值表示符合灰度值 $g = g_0, L = L_0, N = n_0$ 的条件生成概率为下: 其中 $l_0, c_0, n_0$ 分别代表像素的特定灰度值, LBP值, 与NLBP值。

$$P((r, g, b) | L = l_0, C = c_0, N = n_0) \quad (7)$$

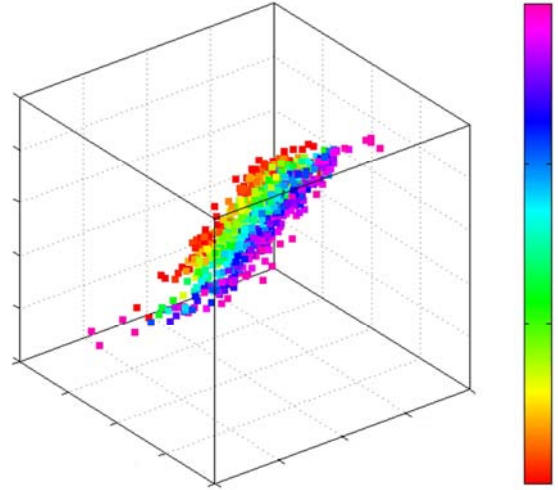


图23 亮度-局部二值特征(LBP)-次局部二值特征(NLBP)-颜色的四维分布4-D。

根据上式表达, R, G, B值会分别满足于分布, 最后再将若干个对应值的其中一个抽取出来做RGB的颜色值。该方法具体实现如下:

1. 对一副灰度图像进行自动着色, 要给出一副灰度图像作为输入;
2. 把含有亮度信息的图像读入, 进行判断. 每一个像素块所对应的LBP值, NLBP值与灰度值。在这个像素点上的亮度值判断是否满足于所训练出来“灰度-局部二值特征-次局部二值特征-颜色”的四维分布。利用公式(7)进行计算RGB值得颜色信息, 即 $f_1(g, l) = R, f_2(g, l) = G, f_3(g, l) = b$ ;
3. 进行图像重构, 分别RGB三通道每一个通道对应一个 $[m \times n]$ 的向量矩阵, 将所计算出的RGB三通道的三个值分别传递到每个矩阵的元素中。
4. 最后将R,G,B三个通道的矩阵值进行合成, 就可以得到一副重塑的目标彩色图像。
5. 该实验具体结果如下:

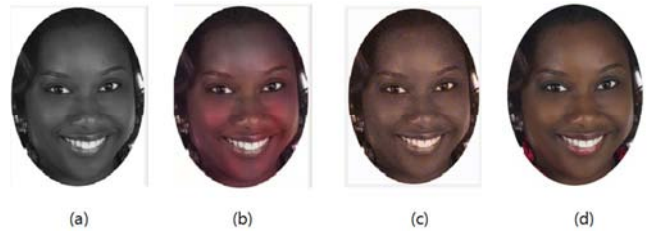


图24 a为灰度图像 b为SatoShi等[5]所采用的卷积网络的结果c为二阶上下文约束的结果 d为彩色原图。

如图24所示, 通过与图24, (b)中[5]研究的对比, 可以发现在照色过程中[5]并没有把最重要的着色部分, 肤色, 成功着色到位。而在图(c)中不难看出, 本文的算法是可以精准求解一幅图像的原始信息, 使其恢复出一个逼真的结果。[5]得研究虽然平滑的将每一个像素值都上色, 但是因为其用卷积网络提取的特征并没有完全的识别所以产生了上色错误的现象。这里本研究做出了一定的取舍, 在嘴唇这一块的特征是特殊的一块, 在其特

征值在灰度相近的情况下很难训练出一种与肤色完全不同的颜色值的分布。通过取舍，可以看出在鼻梁，眼睛，眉毛，头发，脸颊这一块都有清晰的不同的颜色，而且

是可行且生动的。而反观[5]的研究，因为训练的图像庞大，使的此研究虽然对不同自然图像作用广，却耗时耗能且效果不佳，有许多的局限性。



图25 0阶约束，1阶约束与2阶约束方法比较图a为利用0阶上下文所得结果 b为利用一阶上下文约束所得到的结果c为利用二阶上下文约束所得到的结果。



图26 LBP-NLBP 图像对比图左 为彩色原图 中 为 LBP图像 右 为NLBP图像。

在证明了该方法比satoshi等研究对自然人脸图像效果好之后，本文还需要针对是否二阶LBP着实有效这个问题进行探讨。如图25所示，通过用二阶LBP的图像生成并与一阶LBP对比，从大体的着色我们可以看出二阶的LBP所生成的图像相对一阶LBP的图像更加贴近肤色，减少了一些不自然的暖色粒。在平滑度这点上，二阶LBP所生成的图像有明显的优势，整体的着色更加自然，更为真实。从图26可以看到两种LBP模式的略微不同，相对于原LBP图像，NLBP图像的覆盖面积更广，所以局部特征可以完整的显现出来，如NLBP图像就清楚的显示出鼻梁，眼眶等细节。因为二阶LBP算法所考虑的不仅仅是周围的特征信息，还考虑到了更大范围的可能的局部特征，使得在机器学习后基本囊括了各个像的局部特征。通过增加了NLBP的约束，着色结果得到了增强。

### 3. 实验 Experiments

对一副灰色图像进行彩色化后，彩色化的效果如何，需要加以评判。图像的测量方面大致分为两块：图像的逼真度和图像的可理解度。图像的逼真度表示对已处理的彩色化图像与某一原始图像的偏差。在本章3.1中，本文会通过客观的评价标准来衡量并评估本研究逼真度，本文分别会将本文的三个算法的研究与已有研究[5]进行比较，通过计算MSE等用于衡量彩色化算法的值可以清

晰的得到结果。图像的可理解度是指人类对已处理图像的一个理解分析，从主观的判定是否真实来进行测量。在3.2中，本文设计了一个针对于彩色化照片是否自然的调查。本调查作用于广大的群体，从而得出一个主观性的分析。最后3.3部分本文针对黑白老照片也进行了彩色化。通过结合种种分析及实验，本文可以得出一个全面且具体的结果。

#### 3.1. 基于上下文信息算法效果的客观评价

为了量化彩色化图像与原彩色图像的相似程度，本文利用图像均方误差Mean Square Error (MSE) 以及峰值信噪比Peak Signal-to-noise Ratio (PSNR) 作为客观的评价方式来衡量一个彩色化过后的目标图像是否与原图相似。两个的数学公式如下：

$$MSE = \frac{\sum_{c \in \{r,g,b\}} \sum_{0 \leq i \leq M} \sum_{0 \leq j \leq N} (c_{ij} - c'_{ij})^2}{3 \times M \times N} \quad (8)$$

$$PSNR = 10 \times \log_{10} \frac{255 \times 255}{MSE} \quad (9)$$

这里M×N代表着原图像的大小，c代表通过算法所处理出来的r,g,b三通道的值，而c'则代表着原图的r,g,b三通道的值。从计算公式可以看出，两幅照片的像素值差异越大，而MSE则为像素值差异越小，MSE值则越小。用



两个分别不同的衡量指标可以从不同角度来评判，因此是合理的。

试验采用的平台为Intel Iris Plus Graphics 650, 1.5G显存，运行Mac OS high Sierra。本文的时间运行复杂度较低，方法的更新仅是上升一个维度。本研究使用的图像

库均为图像不超过256个像素值的图像，总共图像数约为5K张。参考图像与目标图像的长与宽均不超过600像素，方法1训练时间为10-15分钟，方法2训练时间为25分钟，方法3的训练时间为30分钟。

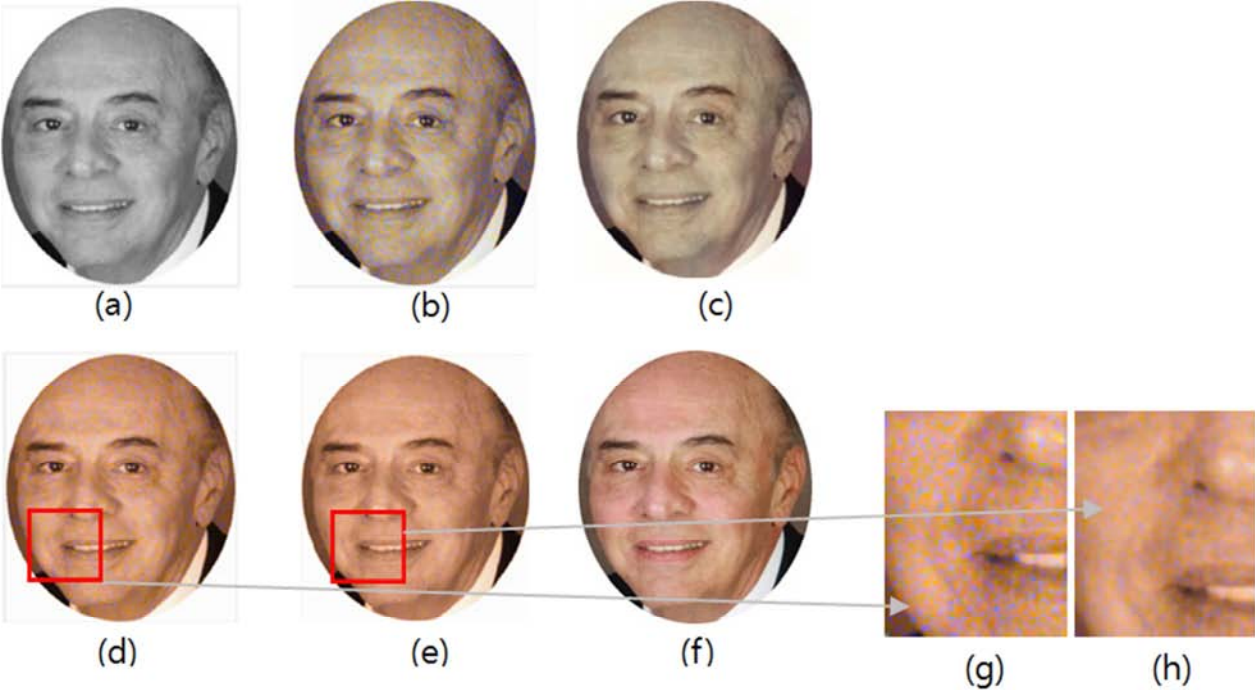


图27 (a) 输入的灰度图像 (b) 单一约束着色结果 (c) [5]等研究结果 (d) 一阶LBP结果 (e) 二阶LBP结果 (f) 彩色原图 (g) 为放大后一阶LBP结果 (h) 为放大后二阶LBP结果。

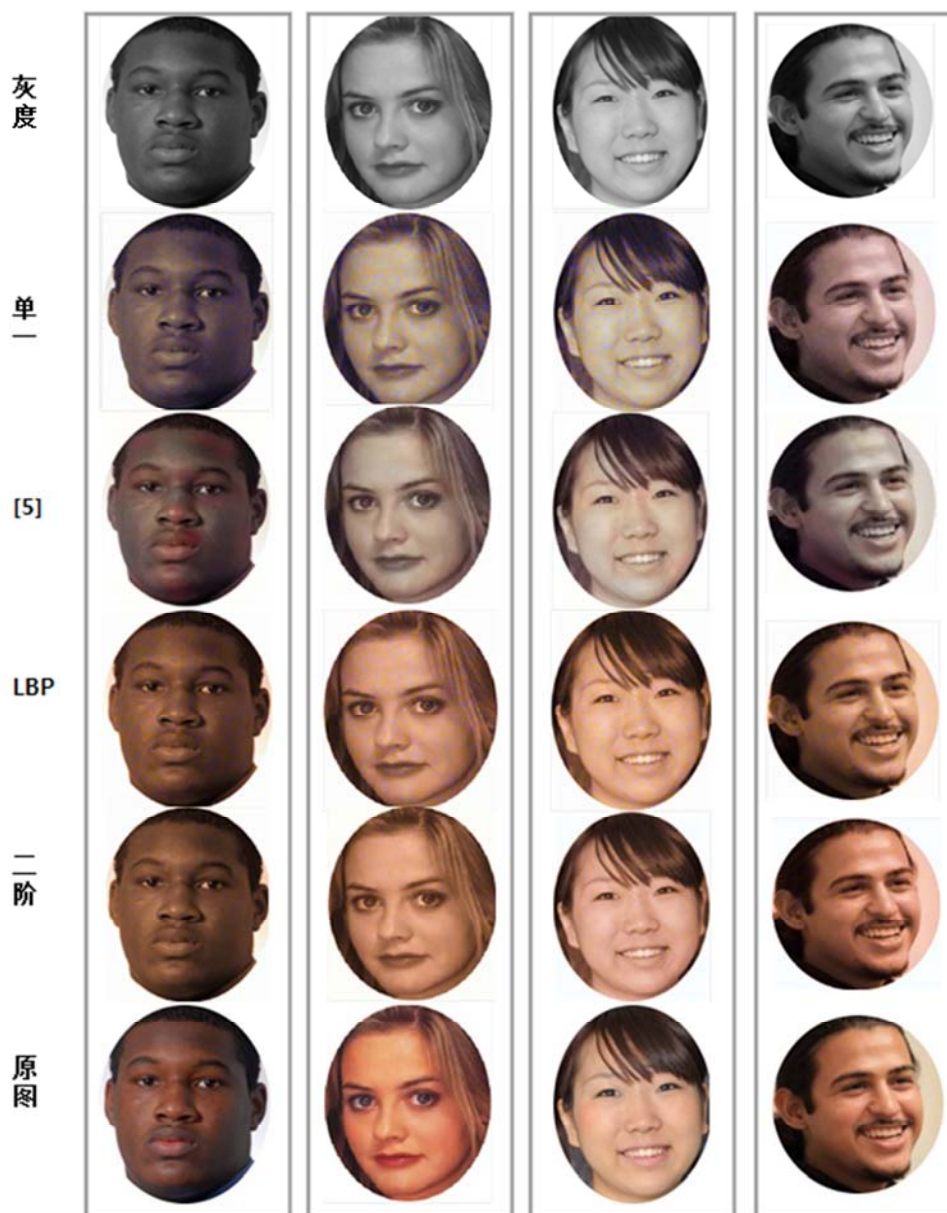
本章首先采用了一组典型的人像作为实验原图，图27展示了本文算法与Satoshi等[5]研究的算法的对比图，表1展示了与图27相对应的客观评价的MSE,PSNR值。通过一个客观的，全面的测评，可以看出基于单一灰度值的彩色算法(图b)没有达到一个好的效果，这因为单一约束求解函数得到可行的值过于广泛，虽然在全局图像是偏白色的肤色，但是在细节的像素粒中可以看到许多噪音值，即上色错误的地方。而[5]得研究就对于平滑度来说比单一约束的彩色算法更优，细节处理也要优于单一约束。但同时，我们可以清晰地看出使用卷积神经网络的算法在着色领域块上色的并不饱满，是一个偏较冷色的颜色值，这是因为在卷积网络所得出来的彩色结果会有一定程度的互相干扰，导致上色出现了很大部分的上色出现差错。而两个本文的基于LBP特征的彩色化算法针对自然着色有明显的提高，现算法将先前的错误的颜色通过特征值的训练，最终获得了饱和度高的着色结果。

而从图g,h可以看出，两幅图像大致着色没有明显的变化，而是在细节上有显著的卓效。这些会使的本文最终研究的算法，二阶LBP获得了一个平滑度更高的着色结果，优化效果明显。这些在根据MSE，PSNR值表中也有相应的体现，本文的算法在MSE上都是远超前于[5]所研究的算法，都是小于0.03的值。表明了颜色差值这一块是占优的。我们还可以发现随着增加的约束（LBP-次LBP）越多，MSE值也随着降低。而在检测图像与原图质量的PSNR值里，[5]的研究是一个平滑度高的图像，相对应的细节也不会流失，所以在与单一约束的PSNR的比较中会有一个PSNR值左右的优势，基于LBP的两个研究要好于[5]，也随着之前的规律，随着约束的增多，图像的质量也有所增加。综上所述，基于“灰度-LBP-NLBP-颜色”的二阶LBP算法针对于自然人脸图像是高效，可行且逼真的。

表1 基于MSE，PSNR对本文演技算法的客观评价。

	单约束算法	[5]Satoshi等算法	一阶LBP	二阶LBP
MSE	0.0256	0.0249	0.0111	0.0091
PSNR	15.9110	16.9497	19.5250	19.5586





**图28** 本文将实验结果作用于一个人脸库中，注意本实验考虑到了不同人像，人种的多样性。从不同种族挑选了典型的人像进行比较，（包括亚裔，非裔，拉丁裔，美裔）实验顺序如前文一样，从上往下，从灰度到[5]到一阶二阶在到原图。

在这部分里，在图28中本文将五张具有代表性的人像进行了着色，并用本研究的分别三个算法与Satoshi[5]等研究以及图像原图进行比较。我们可以发现通过本文自身算法的进阶，图像的质量可以得到明显的改善。随着约束的增加，可以发现四组图像着色明显是朝着原图的方向靠拢的，而且要明显优于[5]的算法。[5]的算法采用的是卷积网络学习的算法，是通过大量的训练然后得到一个经验值来着色，所以此算法可以广泛的利用于不同种类的算法之中。我们可以看出[5]的上色在图组1的肤色是有上错的情况的，这是因为大量的训练也会的导致过多的相匹配的分布，所以导致了上色错误。像本文之前提到的，本文采用的是人像库，所以会使的图像着色过程中更为真实及自然。为了更客观的去测评本次实验，本文还针对三组统计了MSE与PSNR值（MSE所得值越小，

图像颜色越接近。PSNR值越高，图像质量越高），可以看出在MSE上本文的算法有明显的优势，而在PSNR值的对比上平均要比[5]的高出约一个PSNR值以上（基于二阶平均PSNR约为20.4，[5]的平均约为19.5）。从这也证明了本文的算法是可以针对一大类自然图像中利用较小的图像库达到一个比卷积网络更好的效果，同时此算法是高效且实用的。

### 3.2. 真实应用

#### 3.2.1. 与现有手工彩色化方法比较

本文同时测试了旧时代的黑白相片，因为时代不同的原因，现今的照片与过往的照片有巨大的差异。同时这类黑白相片是没有彩色原图的，这也使得利用全自动算法着色成为了可能。因为先前本文训练的是通过将彩

色原图转换过来的“伪”黑白图像，对于这类的图像会比以往着色相对没有经验。如图29所示，本文也将算法作用于一些著名黑白的老照片当中，可以看到本文基于二阶上下文信息的算法在没有训练过了老黑白照片的情况下依然产生出了一个不错的结果。本文将结果对比于瑞

典的艺术家手动上彩的结果进行了比较，可以看出手动上色的结果是非常逼真的，但是其所要耗费的时间通常在15天左右，而反观本文全自动的算法可以在几秒内就生成一个可行的结果，是一个更高效的过程，适用于批量的灰度图像彩色化处理。

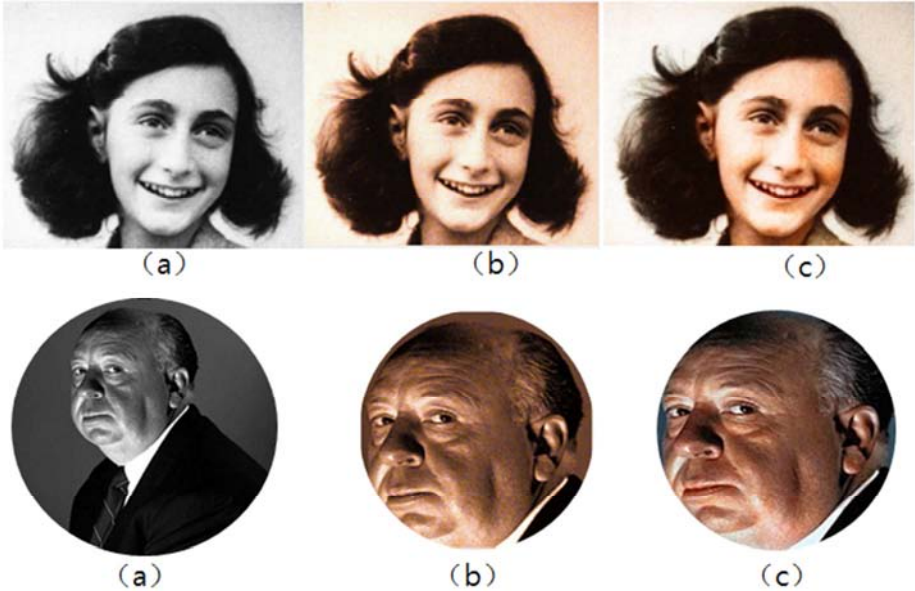


图29 分别采自于著名的“安妮之家”与著名的恐怖连续剧“Alfred Hitchcock”的编辑 (b)为二阶LBP约束的结果 (c)为瑞典艺术家Sanna Dullaway 手工上色的结果[13]。

3.2.2. 与现有研究比较

本文也将算法应用于更广范围的老黑白图像中，同样本文也取了两组的只有黑白的老照片，并与[5]的研究算法进行了比较。如图30所示，第二列所展示的[5]的着色结果，第三列所展示的是本文算法的结果。可以看出，

[5]的颜色重塑结果获得是一个冷色调的结果，跟黑白原图比较相近。而基于二阶上下文信息的算法在着色效果上更为饱满，对应肤色的效果，证明出在黑白人脸的着色中本文的算法是有效的。

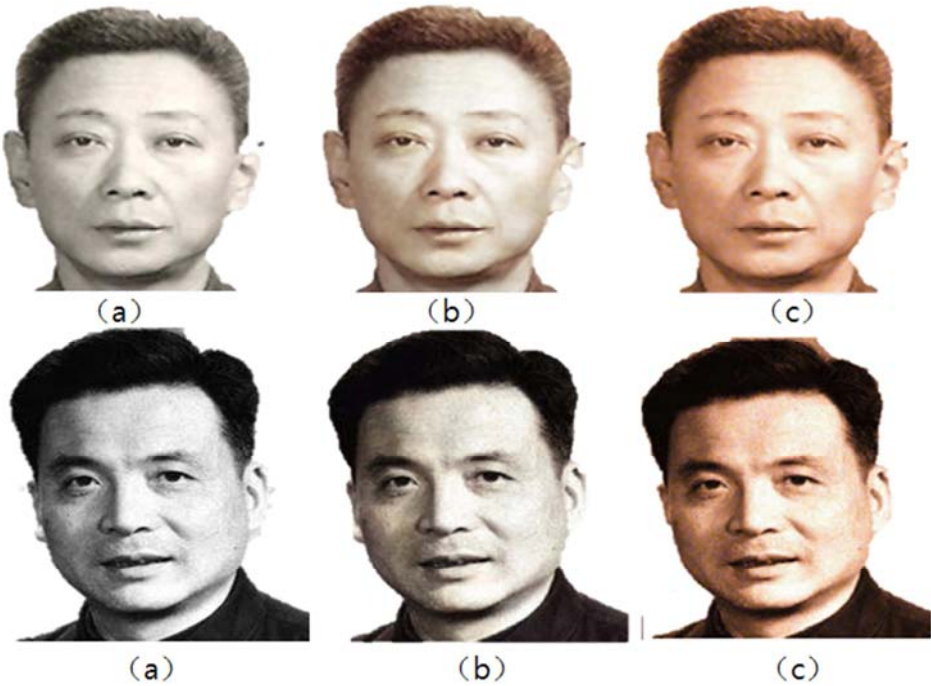


图30 a 为两组图分别采自于采自于80年代的胶卷相片，“两弹一星”的创始人，邓稼先的黑白胶卷相片 b 为[5]的结果 c为二阶LBP约束的结果。

## 4. 总结及展望 Conclusions & Expectations

### 4.1. 工作总结

本文通过对图像处理的深入研究，将基于局部二值模式的特征应用在图像颜色重构领域，提出了分别三个递进的基于上下文信息的彩色算法。首先依据数据库的彩色图像训练得到“亮度-LBP-NLBP-颜色”的四维分布；然后利用灰度，LBP纹理特征和NLBP纹理特征的约束性求解目标图像块的RGB值，对像素进行着色。通过实验证明了这个算法在人脸图像中可以减少上色错误的可能，重塑出一个平滑度，饱满度高的彩色图像。

此研究的亮点在于：

- 1) 提出了一个新的训练架构，添加新颖的约束项，使其在图像中效果极佳
- 2) 本研究在一个很小的数据库中，训练出算法的彩色结果要明显优于以往的基于卷积神经网络自动化彩色算法；
- 3) 本文提出的着色方法是完全自动的，过程中不需要人为干涉；
- 4) 推动彩色化的普遍化，使得老年代的黑白照片颜色还原更加有可操作性，有望促进黑白图片的批量彩色化；
- 5) 算法的训练与生成过程是迅速，高效的。

### 4.2. 未来展望

在目前的研究上，本文的算法还存在着一定的局限性。

对于基于二阶上下文信息的灰度图像着色算法而言：

- 1) 首先，该方法对于纹理特征不是很明显的区域着色效果不佳，如人脸的嘴唇这一块，这是因为本文在学习分布的过程中是考虑到邻域的灰度值的信息的，因为嘴唇的灰度值与不同肤色的人脸的灰度值是非常相近的，使的在求解嘴唇这一部分的颜色值的时候会产生上色错误。在以后的工作中可以针对于嘴唇这一块做一个训练，以真正的使人脸彻底着色正确。
- 2) 其次，该算法虽然对人脸完成了好的自动着色效果，但是针对其他种类的图像着色效果不佳。这是因为本文训练的图像库是只针对于人脸的图像库，在今后的工作可以针对每一大类的图片分别使用使用不同的图像库训练出分布，使的本算法的着色面更广。

但在未来的研究上，本文的算法可以在以下几方面进行延续：

- 1) 因为本文的在训练与生成的速度都相当快，所以可以添加一个时间的维度，使得可以利用本算法实现从视频上的彩色化。
- 2) 同时为了使得广大的群体可以使用本文的模型，可以制作一个云端的模型以网站的形式使得大众可以轻易使用。

## 致谢

在论文写作之中，我有幸得到了中山大学数学系冯国灿教授的悉心指导。感谢冯老师在研究阶段对我提供的帮助与鼓励，让我顺利的完成研究。

## 参考文献

- [1] Levin, A., Lischinski, D., & Weiss, Y. (2004, August). Colorization using optimization. In *ACM transactions on graphics (tog)* (Vol. 23, No. 3, pp. 689-694). ACM.
- [2] Di Blasi, G., & Reforgiato, D. (2003). Fast colorization of gray images. *Eurographics Italian*.
- [3] Luan, Q., Wen, F., Cohen-Or, D., Liang, L., Xu, Y. Q., & Shum, H. Y. (2007, June). Natural image colorization. In *Proceedings of the 18th Eurographics conference on Rendering Techniques* (pp. 309-320). Eurographics Association.
- [4] Zhang, R., Isola, P., & Efros, A. A. (2016, October). Colorful image colorization. In *European Conference on Computer Vision* (pp. 649-666). Springer, Cham.
- [5] Iizuka, S., Simo-Serra, E., & Ishikawa, H. (2016). Let there be color!: joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification. *ACM Transactions on Graphics (TOG)*, 35(4), 110.
- [6] Deshpande, A., Rock, J., & Forsyth, D. (2015). Learning large-scale automatic image colorization. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 567-575).
- [7] Liu, S., & Zhang, X. (2012). Automatic grayscale image colorization using histogram regression(Tech.). Tianjin University. doi:1673-1681
- [8] Kumar, T., & Verma, K. (2010). A Theory Based on Conversion of RGB image to Gray image. *International Journal of Computer Applications*, 7 (2), 7-10.
- [9] Yan, D., Bai, L., Zhang, Y., & Han, J. (2018). Multi-sparse dictionary colorization algorithm based on the feature classification and detail enhancement. *Optical Review*, 25(1), 78-93.
- [10] Hao, K. (2012). 基于稀疏编码的图像视觉特征提取及应用 (Y2060075, Tech.). The Hebrew University of Jerusalem. 硕士毕业论文
- [11] Lloyd, J. (2017, May 05). The Process. Retrieved from <http://dynamichrome.com/process/>, 2018/7/25
- [12] K. (n.d.). Famous Black and White Photos Restored in Color. Retrieved from <https://www.amusingplanet.com/2012/05/famous-black-and-white-photos-restored.html#modal-one> ,2018/8/3



- [13] Welsh, Tomihisa, Michael Ashikhmin, and Klaus Mueller. "Transferring color to greyscale images." *ACM Transactions on Graphics (TOG)*. Vol. 21. No. 3. ACM, 2002.
- [14] Pietikäinen, Matti. "Local binary patterns." *Scholarpedia* 5.3 (2010): 9775.
- [15] Lloyd, Jordan. "The Hoover Dam Under Construction, 1935." *Medium*, Dynamichrome Viewfinder, 5 Apr. 2016, <https://medium.com/dynamichrome-viewfinder/the-hoover-dam-under-construction-1935-2148d2ba26d3>.