一种利用回归神经网络黑白图像着色算法研究

余梓唐
（义乌工商职业技术学院 机电信息分院，浙江 义乌 322000）

摘　要： 研究了利用神经网络对序列黑白（灰度）图像进行着色的问题。针对以往基于人工或者半自动化技术的黑白图像着色技术效率低下、视觉效果较差的缺陷，提出了一种利用三层神经网络、无须人工干预的图像自动着色算法。首先将灰度图像分割成小块，通过对小块提取灰度特征、空间特征等作为神经网络的输入，训练得到一个回归神经网络。在着色过程中，可以利用该神经网络将图像中各像素由灰度空间投影到一个经过压缩的色彩空间，从而实现了图像的自动着色过程。实验结果显示本方法能有效地将灰度图像着色，并且由于使用了一个压缩的色彩空间，使得计算效率和着色效果都得到了有效的提高，能很好地逼近原始的真实图像。

关键词： 图像着色；图像序列；神经网络；灰度空间
中图分类号： TP391　　**文献标志码：** A　　**文章编号：** 1001-3695(2012)04-1595-03
doi:10.3969/j.issn.1001-3695.2012.04.111

Neural networks based colorization for black-white images

YU Zi-tang
(*School of Electro-Mechanical & Information Technology*, *Yiwu Industrial & Commercial College*, *Yiwu Zhejiang* 322000, *China*)

Abstract: This paper presented a neural network based colorization algorithm which could transform gray-scale images into color images. The traditional colorization methods have several shortages such as consuming time and performing is bad. First, the method segmented gray-scale image into smaller blocks. Then, based on a neural network which took the gray-scale and the location features of each block as its input, it could map the original gray-scale image into a compressed color space to reconstruct a new color image. This method had several advantages. First, the used features were simple. Second, because it used the compressed color space, the efficiency of the method was very high. The experimental results show that the result images are very close to the original color image, and this algorshm can provide better visual performance than others.

Key words: colorization; image sequence; neural network; gray space

0　引言

从 20 世纪 80 年代开始，在图像处理和电影电视产业领域，出现了利用着色技术对传统黑白照片或影片进行自动着色的技术，从而使得传统的黑白照片或影片能够提供更为丰富的色彩信息，满足观赏者的视觉需求。这一技术最初利用人工对图像或影片中的各个像素进行颜色的标定和着色，其缺点是效率低下。随着近年来数字技术和计算机技术的迅速发展，图像自动着色技术得到了长足的发展，并在其他图像处理领域得到了广泛的应用。例如在医学图像处理中，利用着色技术对人体各器官进行自动着色，便能有效地区分人体各个正常器官与病变器官。而对影视作品的自动着色技术，则使得观众在观赏老作品的时候能获得新鲜的视觉感受[1]。

对于各种自动着色技术，其主要思想是利用色彩空间中的色彩向量来替换各个像素原始的灰度或亮度值。通常所采用的色彩空间为 RGB 色彩空间，即在三维色彩空间中，三个坐标分别代表红绿蓝色的强度。由于从色彩空间到灰度空间变换的不可逆性，要从灰度空间返回到色彩空间通常需要有一定的人工智能的帮助，即在对黑白或灰度图像进行自动着色之前，需要部分的人工参与来判定着色是否正确，并对后续的自动着色进行调整。

在以往的自动着色技术中，Wilson 提出的着色算法是最早的尝试，主要使用在早期的黑白电影着色中，所能恢复的色彩较为简单[2]。从那以后，许多人工或半自动的图像着色技术被提出来。在多数方法中，一幅黑白图像被切分成若干小区域，然后人工或依照参照的彩色图像对各个小区进行颜色标定，从而达到着色的目的；再将这幅已经被着色的图像作为参照，对后续的图像进行着色标定。从这一过程可以看出，虽然将现代计算机技术和电子技术运用到了图像着色之中，传统的着色方法依然费时费力。

因此在本文中，利用现代人工智能的方法，即神经网络学习算法，通过对部分黑白图像着色过程和结果进行训练与学习，从而实现图像着色的自动化操作。虽然基于机器学习的着色算法需要经过较长时间的数据收集和训练，但是其着色过程效率得到极大的提高，在后期处理中能节省大量的人力物力，并能显著提高着色的准确率和视觉效果。

收稿日期： 2011-08-07；**修回日期：** 2011-09-14
作者简介： 余梓唐（1966-），男，浙江义乌人，讲师，硕士，主要研究方向为智能计算、模式识别等（yztyw@ywu.cn）.

1 自动着色算法基本框架

本文提出的黑白(灰度)图像自动着色算法主要包括两个部分:a)寻找与当前黑白图像最为相关的彩色图像作为参考图像;b)利用参考彩色图像、黑白图像自身所提取的特征以及神经网络技术,实现对黑白图像的着色。

该系统中最为核心的部分是对人工神经网络的训练和使用。在训练之前,必须人工建立训练数据库,并且标定出目标黑白—彩色图像对作为训练系统的输入和输出。

在以往的方法中,常采用的是将一个影像序列中的第一幅图像进行人工着色标定,作为后续着色的参考图像。但是这样就不能完全实现图像自动着色;并且当图像序列中的场景发生变化后,又需要人工重新进行标定与着色,使得人的负担加重。因此,本文算法将建立一个标准的彩色图库,自动搜寻与黑白图像相关度最大的彩色图像作为着色的参照图像。

本文算法的核心是开发一种基于人工智能技术的着色算法,通过对大量的模式识别与机器学习算法的筛选,最终选定具有多层感知器的人工神经网络技术作为识别分类器。这一分类器已在模式识别领域得到了广泛的使用和研究,其多层神经网络构成了一个非线性的投影变换,能将一组低维特征投影到一个高维空间,从而实现从低维的黑白图像投影得到高维彩色图像的目的。

为了构建神经网络的训练图库,取出一个图像序列中的第一幅图像对其进行颜色标定,然后以这幅图像作为训练图像来对神经网络进行训练学习。

1.1 神经网络的构建

针对本文所提出的着色问题,在训练的时候,神经网络的输入为训练图像中单个图像的像素灰度。为了提供各个像素的空间信息,本文在输入中增加了各个像素的 x 坐标和 y 坐标,而输出则为三个颜色通道中各个颜色的强度。

本文所设计的神经网络有三层,分别为输入层、中间层和输出层。其中,输入层神经元个数为 3,即输入像素的灰度值及其相应的坐标;中间层神经元个数为 20,神经元所采用的相应函数为 sigmoid 方程;而输出层则含有 3 个神经元。输入层和输出层的神经元均采用线性函数作为其相应函数。图 1 为隐含层(中间层)神经元采用的响应函数表达式及其曲线。

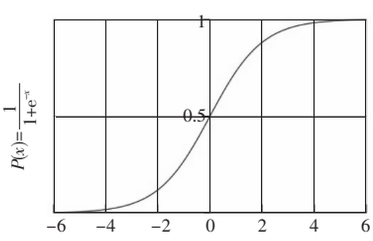

图 1 sigmoid 函数曲线

对于三个神经网络输入特征的选择,主要依据是相邻两帧图像之间,在相同像素附近的颜色具有较大的相关性。如果考虑到相邻像素间的相关性,可以更多地提取与像素局部邻域有关的纹理特征作为神经网络的输入,如文献[3]中所提及的 Laws 滤波器。该滤波器集合包含一组共九个滤波器,其基本原理是分析某个像素邻域内的颜色相关性,因此也可以利用该组滤波器的输出作为神经网络的输入特征。由此构成的神经网络的输入层因此具有 11 个输入神经元。最终的神经网络结构如图 2 所示。

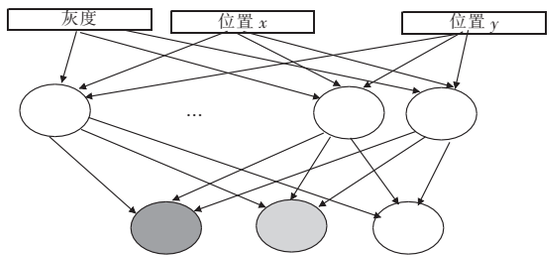

图 2 神经网络结构

1.2 特征的提取与色彩空间的压缩

为了减少计算量,并利用图像中的纹理信息,在提取图像各个像素特征之前,将图像分割为 5×5 像素的小块,每个小块构成一个超像素,其对应的灰度均值、色彩均值及滤波特征分别作为神经网络的输入,并且超像素中的 25 个像素共用一个坐标(x,y)。

由于图像着色的过程可以看成是一个从 256 级灰度空间到 $2^{24}=16$ M RGB 色彩空间的投影过程,如果能将目标色彩空间进行某些限定,即压缩目标空间的颜色数量,则能有效地提高神经网络的训练效率和识别效率。因此,在保证提供良好人眼视觉效果的同时,将色彩空间压缩到 $2^{12}=4\ 096$ 种 RGB 色彩,作为最终神经网络输出结果的色彩空间。

1.3 性能的测定

为了测定图像着色系统的性能,使用不同颜色通道的均方差来衡量两幅图像之间的颜色相关度。假设有 A 和 B 两幅图像,其计算过程如下:

$$\mathrm{MSE}_R=\frac{1}{h\times w}\sum_{i=1}^{h}\sum_{j=1}^{w}(R_{i,j}^{\mathrm{A}}-R_{i,j}^{\mathrm{B}})^2 \tag{1}$$

$$\mathrm{MSE}_G=\frac{1}{h\times w}\sum_{i=1}^{h}\sum_{j=1}^{w}(G_{i,j}^{\mathrm{A}}-G_{i,j}^{\mathrm{B}})^2 \tag{2}$$

$$\mathrm{MSE}_B=\frac{1}{h\times w}\sum_{i=1}^{h}\sum_{j=1}^{w}(B_{i,j}^{\mathrm{A}}-B_{i,j}^{\mathrm{B}})^2 \tag{3}$$

其中:h 和 w 分别为图像的高和宽;R、G、B 分别指代各个不同颜色通道的强度。对于所有图像之间的颜色相关度,采用如下公式计算:

$$\mathrm{MSE}_R=\frac{1}{h\times w\times nf}\sum_{k=1}^{nf}\Big[\sum_{i=1}^{h}\sum_{j=1}^{w}(R_{i,j}^{f1}-R_{i,j}^{f2})^2\Big]_k \tag{4}$$

$$\mathrm{MSE}_G=\frac{1}{h\times w\times nf}\sum_{k=1}^{nf}\Big[\sum_{i=1}^{h}\sum_{j=1}^{w}(G_{i,j}^{f1}-G_{i,j}^{f2})^2\Big]_k \tag{5}$$

$$\mathrm{MSE}_B=\frac{1}{h\times w\times nf}\sum_{k=1}^{nf}\Big[\sum_{i=1}^{h}\sum_{j=1}^{w}(B_{i,j}^{f1}-B_{i,j}^{f2})^2\Big]_k \tag{6}$$

$$\mathrm{MSE}_{\mathrm{total}}=\frac{\mathrm{MSE}_R+\mathrm{MSE}_G+\mathrm{MSE}_B}{3} \tag{7}$$

其中:nf 表示图像帧的总数,f_1 和 f_2 表示相对应的图像帧。

为了计算相对于 256 级灰度的不同色彩强度的变化率,对最终的相关度进行归一化处理:

$$\mathrm{MSE}_{\mathrm{total}}\%=\frac{\mathrm{MSE}_{\mathrm{total}}}{256}\times 100 \tag{8}$$

最终,计算着色系统的性能用如下公式:

$$P=\frac{\mathrm{MSE}_{\mathrm{total}}^{\mathrm{BW-C}}-\mathrm{MSE}_{\mathrm{total}}^{\mathrm{C-Cz}}}{\mathrm{MSE}_{\mathrm{total}}^{\mathrm{BW-C}}}\times 100 \tag{9}$$

其中:MSE_{total}^{BW-C}表示黑白图像与原始彩色图像之间的均方差,MSE_{total}^{C-Cz}则表示通过着色后的图像与原始彩色图像之间的均方差。

2　实验设置和结果

在本文中,采用 AVI 格式的视频图像进行训练和测试,其分辨率为 640 ×480。图 3 显示了一组从视频中截取的原始图像和相应的灰度图像示例。

(a) 原始图像　(b) 灰度图像　(c) 原始图像　(d) 灰度图像

图 3　实验图像样本

实验设置如下:首先从视频中采集一组彩色序列图像;然后通过对各个颜色通道的强度加权平均,将彩色图像转换成为相应的灰度图像;再采用本文提出的着色方法对灰度图像进行着色操作,最终得到着色化以后的彩色图像。

在图 4 和 5 中显示了图 3 所示样本序列中不同灰度图像的着色结果。

(a) 第 10 帧　(b) 第 30 帧　(c) 第 45 帧　(d) 第 70 帧

图 4　飞机序列图像着色结果

(a) 第 5 帧　(b) 第 20 帧　(c) 第 40 帧　(d) 第 60 帧

图 5　湖面序列图像着色结果

从实验结果可以看到,由于采用了色彩空间压缩,经过着色处理的图像颜色没有原始图像的色彩丰富,但着色后的图像在视觉上依然能够提供较好的色彩观感效果。

为了对比不同特征和神经网络对于图像着色的影响,本文进行了不同的测试,并将结果显示在表 1 中。

从以上实验结果可以看到,使用各个像素的灰度特征以及坐标特征并结合色彩空间压缩的方法,能有效地提高图像着色的准确率。从表 1 中可以得到如下结论:在使用合适的色彩空间压缩算法后,不但能有效地减少系统的计算量和过高特征维数给神经网络的训练及识别造成的不稳定性因素,还提高了系统的着色效果,使得图像更接近于原始的彩色图像。

表 1　图像着色实验结果

实验号	实验设置	飞机序列/%	湖面序列/%
1	仅使用灰度	37.3	20.5
2	实验 1 + 坐标	62.0	55.7
3	Laws 滤波	560	47.5
4	实验 1 + 色彩空间压缩	91.1	79.9
5	实验 2 + 色彩空间压缩	93.2	83.0
6	实验 3 + 色彩空间压缩	92.5	80.1

3　结束语

本文提出了一种基于神经网络的灰度(黑白)图像着色方法。通过利用像素的灰度信息以及空间信息,以一个相关图像序列中的第一幅图像作为训练图像,能快速地训练得到一个神经网络,并将灰度图像投影到一个压缩的色彩空间中,形成一幅彩色图像。实验结果显示,本文的算法能提供具有良好视觉效果的着色结果。利用其他类型的特征或色彩空间表达,如 YIQ 色彩空间等,本文所提出的方法构架能获得更好的实验效果。

参考文献:

[1] WELSH T, ASHIKHMIN M, MUELLER K. Transferring color to geyscale images[J]. ACM Trans on Graphics, 2002, 21(3): 277-280.

[2] IRONY R, COHEN-OR D, LISCHINSKI D. Colorization by example [C]//Proc of Eurographics Symposium on Rendering. 2005: 201-210.

[3] MONADJEMI A. Towards efficient texture classification and abnormality detection[D]. Bristolvk: Dept of Comptuter Seience, University of Bristol, 2004: 19-21.

[4] 廖飞雄,马良. 图着色问题的启发式搜索蚂蚁算法[J]. 计算机工程,2007,33(16): 191-192,195.

[5] 朱黎博,孙韶媛,谷小婧,等. 基于色彩传递与扩展的图像着色算法[J]. 中国图象图形学报,2010,15(2):200-205.

[6] 赵国英,李华. 人体脸部灰度图像上色的改进算法[J]. 计算机辅助设计与图形学学报,2004,16(8):1051-1056.

[7] QU Ying-ge, WONG T T, HENG P A. Manga colorization[J]. ACM Trans on Graphics, 2006, 25(3): 1214-1220.

[8] YATIZV L, SAPTRO G. Fast image and video colorization using chrominance blending[J]. IEEE Trans on Image Processing, 2006, 15(5):1120-1129.

(上接第 1574 页)

[4] 范九伦,赵凤. 灰度图像的二维 Otsu 曲线阈值分割法[J]. 电子学报,2007,35(4):751-755.

[5] MEDINA-CAMICER R, MADRID-CUEVAS F J. Unimodal thresholding for edge detection[J]. Pattern Recognition, 2008, 41(7): 2337-2346.

[6] 冯斌,王璋,孙俊. 基于改进粒子群算法的二维阈值图像分割[J]. 计算机应用研究,2008,25(8):2402-2404.

[7] 江禹生,宋香丽,任晶晶. 基于遗传算法的二维 Otsu 算法改进[J]. 计算机应用研究,2010,27(3):1189-1191.

[8] 范九伦,赵凤,张雪峰. 三维 Otsu 阈值分割方法的递推算法[J]. 电子学报,2007,35(7):1398-1402.

[9] WANG Na, LI Xia, CHEN Xiao-hong. Fast three-dimensional Otsu thresholding with shuffled frog-leaping algorithm [J]. Pattern Recognition Letters, 2010, 31(13):1809-1815.

[10] REALE M, HUNG T, YIN Li-jun. Pointing with the eyes: gaze estimation using a static/active camera system and 3D iris disk model [C]//Proc of IEEE International Conference on Multimedia and Expo. [S.l.]: IEEE, 2010: 280-285.

[11] WANG Lu, YANG Gong-ping, YIN Yi-long. Fast iris localization based on improved Hough transform[C]//Proc of the 5th International Conference on Rough set and Knowledge Technology. Berlin: Springer-Verlag, 2010: 439-446.