基于 *k*–Means 算法的彩色 QR 码识别

doi:10.3969/j.issn.1674–7100.2017.05.010

贾　丹[1]　尤　飞[1]
张庆立[1]　曾志高[2]

1. 上海理工大学
出版印刷与艺术设计学院
上海　200093
2. 湖南工业大学
计算机学院
湖南　株洲　412007

摘　要：为了降低彩色 QR 码解码过程中出现的混叠效应，提高彩色 QR 码解码的正确率，提出一种基于 HSV 颜色模型的 *k*–Means 聚类算法。为了选择适合彩色 QR 码的颜色空间模型，通过实验验证了在 RGB, Lab, HSV 3 个颜色模型下 *k*–Means 聚类算法的效果。在 HSV 颜色模型下，根据等欧氏距离的原则建立彩色编码模块的配色模型，最大程度地减小解码中颜色的混叠效应。彩色 QR 码解码预处理阶段，利用基于 HSV 颜色模型的光线补偿的 *k*–Means 聚类算法对彩色编码模块进行颜色分离，以提高解码的精度。研究结果表明：在 HSV 颜色模型下，*k*–Means 聚类效果最好，图像区域分类效果最清晰；所建立的配色模型可以最优地为彩色编码模块配色；基于 HSV 颜色模型的光线补偿的 *k*–Means 聚类算法可以提高彩色 QR 码解码的正确率。因此，建立合理的配色模型进行彩色编码模块的颜色设置，同时采用基于 HSV 颜色模型的光线补偿的 *k*–Means 聚类算法进行颜色分割，可以大幅度地降低彩色 QR 码编码模块之间的混叠效应，从而显著提高彩色 QR 码解码的正确率。

关键词：彩色 QR 码；*k*–Means 聚类算法；HSV 颜色模型；彩色 QR 码的配色模型

中图分类号：TP391.44　　**文献标志码：**A

文章编号：1674–7100(2017)05–0062–07

1　研究背景

黑白 QR（quick response）码具有存储容量大、纠错能力强等优点，因而被广泛应用于交通运输、产品溯源、电子商务等方面[1]。由于二维码应用领域的迅速扩张，黑白 QR 码已经难以满足爆炸式信息传递的时代需求，因此国内外众多学者将研究的方向聚焦到了彩色 QR 码技术上。

彩色 QR 码是建立在黑白二维条码的基础上，即将发展起来的一种新型的大容量的彩色条码技术。彩色 QR 码代替黑白 QR 码，极大地提高了相同版本、相同密度下条码的存储容量。由于该技术在信息存储容量、图形的可塑性、防伪、外观等方面明显优于黑白 QR 码，所以成为国内外学者研究的重点[2]。但是，在其应用和推广方面出现了瓶颈而停滞不前，其重要的原因是，彩色 QR 码在解码过程中，彩色编码模块出现色彩偏离、颜色混叠效应，产生了颜色识别上的困难。

国外彩色条码技术对于颜色识别方面的研究起步较早。HAN T. D. 等[3]首次提出了彩色二维码（color code）的技术，利用参考颜色克服了颜色保真度的问题，为区分每种颜色提供了标准的颜色参考，为彩色

收稿日期：2017–07–17
基金项目：上海理工大学博士启动基金资助项目（1D–13–309–007），模式识别国家重点实验室开放课题基金资助项目（201700009），湖南省教育厅基金资助项目（17C0481）
作者简介：贾　丹（1992–），女，山东省烟台人，上海理工大学硕士生，主要研究方向为图文信息处理，E-mail：jiadan106@126.com
通信作者：尤　飞（1973–），男，安徽全椒人，上海理工大学副教授，博士后，主要从事 CAD/CAM，印刷包装设备自动控制，图文信息处理方面的研究，E-mail：yfreeman@sina.com

二维码技术的发展奠定了基础。N. Bhardwaj 等[4]考虑到移动设备中，颜色都是通过 RGB 颜色表示的，提出了一种基于 Android 系统的彩色 QR 码的解码方法，同时设定了 1 个为用于不同版本 QR 码解码成功的最佳分辨率。M. Kikuchi 等[5]提出了一种新的彩色 QR 码，为了增加编码信息的传输能力，将彩色 QR 码在 YCbCr 颜色模型下分成 3 个通道进行解码，Y 通道采用传统标准 QR 码解码器，Cb 和 Cr 通道则分别采用 1 个专门的解码器。

国内学者基于 QR 码直接扩容技术，对彩色二维码进行了相关的研究。陈元枝等[6]提出基于 Zxing 库函数的彩色 QR 码生成与识别方法，不仅扩充了数据容量，而且保留了黑白 QR 码的可靠性和鲁棒性，但是本算法并不支持扫描纸质档的彩色 QR 码。郑河荣等[7]提出了具有高压缩比汉字编码能力的彩色二维码，按 RGB 模型分通道的操作和图形定位校正，实现解码过程。牛万红等[8]提出了一种彩色 QR 码的设计原理及编解码实现技术，通过扩大码字符号的尺寸，减少其数量，降低算法的空间复杂度。

以上研究成果已经得到了一定的应用，对彩色 QR 码技术的发展提供了理论支撑，但仍存在一些问题。现有的彩色 QR 码解码不用于印刷或者多采用比较精密的解码设备，彩色模块之间产生的混叠效应，使得识别困难，解码过程复杂。本文提出一种基于 *k*-Means 算法的彩色 QR 码颜色分离处理方法，在 HSV（hue, saturation, value）颜色模型下进行彩色编码模块的配色，建立彩色 QR 码编码模块配色模型；在彩色 QR 码解码预处理过程中，采用 *k*-Means 算法进行颜色分割。通过此技术手段，能简单有效地降低彩色编码模块之间的混叠效应，实现颜色模块的分离，从而提高彩色 QR 码解码的正确率。

2　彩色 QR 码识别的关键技术

在彩色图像的处理过程中，颜色模型的选取，在很大程度上关系着彩色图像处理算法的可行性和效果；在彩色图像边缘检测、图像增强和颜色聚类分析时起着重要的作用。本文主要对符合人眼视觉的 Lab 颜色模型和 HSV 颜色模型进行分析。

2.1　颜色模型的特征属性

颜色模型是指使用一组数值来表示颜色的一种抽象的数学模型。颜色模型中的颜色值通常用代表 3 个参数的 3 维坐标来描述，其颜色取决于所使用的坐标。从颜色感知的角度来说，颜色模型可以分成 3 类，如表 1 所示。

表 1　颜色模型分类及特征举例

Table 1　Classification of color space and examples of features

类　型	特　征	举　例
混合型	按 3 种基色的比例合成颜色	RGB, CMYK 和 XYZ 等
非线性亮度 / 色度型	用 1 个分量表示非色彩的感知，用 2 个独立的分量表示色彩的感知	Lab, Luv, YUV 和 YIQ 等
强度 / 饱和度 / 色调型	用饱和度和色度描述色彩的感知	HSV, HSI, HSL 和 LCH 等

1）RGB 颜色模型。根据三基色原理，用基色光单位来表示光的量，则在 RGB 颜色模型中，任意色光 F 都可以用不同分量的 *R*, *G*, *B* 三色相混合而成。调整三色系数 *R*, *G*, *B* 中的任一系数都会改变 *F* 的坐标值，即改变了 *F* 的色值。这一体制并不适应人的视觉特点。因而，产生了其他不同的颜色模型表示法[9]。

2）Lab 颜色模型。与 RGB 颜色模型相比，Lab 颜色模型是一种不常用的颜色模型。Lab 颜色模型由国际照明委员会于 1976 年制定，是一种与设备无关的颜色系统。Lab 颜色模型中的 *L* 分量表示像素的亮度，取值范围是 [0, 100]，即从纯黑到纯白；*a* 表示从红色到绿色，取值范围是 [127, −128]；*b* 表示从黄色到蓝色，取值范围是 [127, −128]。在 Adobe Photoshop 图像处理软件、TIFF 格式文件和 PDF 文档中都可见到 Lab 颜色模型的身影[10]。

从 RGB 颜色模型到 Lab 颜色模型的转换，只能使用 CIE XYZ 颜色模型间接转换，转换公式如下：

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \frac{1}{255}\begin{pmatrix} 0.433\,910 & 0.376\,220 & 0.189\,860 \\ 0.212\,649 & 0.715\,169 & 0.072\,182 \\ 0.017\,756 & 0.109\,478 & 0.872\,915 \end{pmatrix}\begin{pmatrix} R \\ G \\ B \end{pmatrix}, \tag{1}$$

$$\begin{cases} L = \begin{cases} 116\left(Y/(3Y_n)\right)-16, & Y/Y_n > 0.008\,856; \\ 903.3\left(Y/Y_n\right), & Y/Y_n \leqslant 0.008\,856; \end{cases} \\ a = 500\left(f\left(X/X_n\right)-f\left(Y/Y_n\right)\right); \\ b = 200\left(f\left(Y/Y_n\right)-f\left(Z/Z_n\right)\right)。 \end{cases} \tag{2}$$

式（1）~（2）中：*R*, *G*, *B* 为像素的 3 个通道，取值范围均为 [0, 255]；

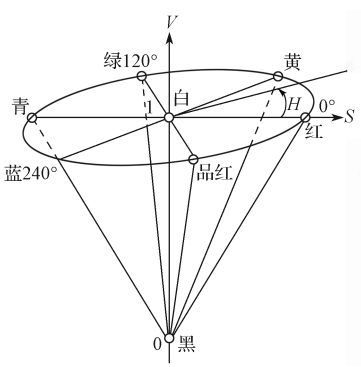
$f(t)=\begin{cases} t^{1/3}, & t>0.008\,856; \\ 7.787t+16/116, & t\leqslant 0.008\,856。 \end{cases}$ 其中 *t* 代表 X/X_n, Y/Y_n 或 Z/Z_n，X_n, Y_n, Z_n 的默认值分别为 95.047，100.000，108.883。

3）HSV颜色模型。HSV颜色模型将颜色的色调、饱和度和明度3个属性相分离、彼此独立，更符合人眼对色彩的感知；同时，HSV颜色模型在光照不足的情况下色彩也表现相当稳定。所以，HSV颜色模型更有利于彩色QR码图像的识别[11]。HSV颜色模型采用色调、饱和度和明度3个属性来表征颜色，如图1所示。

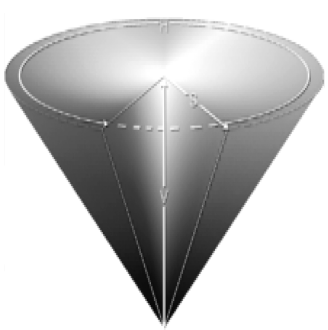
色调 H：用角度度量，取值范围为0°~360°，从红色开始按逆时针方向计算，红色为0°，绿色为120°，蓝色为240°。它们的补色是：黄色为60°，青色为180°，品红色为300°。

饱和度 S：取值范围为0~1，取值越大，颜色越饱和。

明度 V：表示色彩的明亮程度，取值范围为0~1。



a）模型三维表示



b）模型色域

图1 HSV颜色模型

Fig. 1 HSV color space model

RGB颜色模型到HSV颜色模型的转换公式为：

$$\begin{cases} V=\dfrac{1}{\sqrt{3}}(R+G+B); \\ S=\dfrac{\sqrt{3}}{V}\min(R,\ G,\ B); \\ H=\begin{cases}\theta, & B\leqslant G; \\ 2\pi-\theta, & B>G。\end{cases} \end{cases} \tag{3}$$

式中 $\theta=\arccos\dfrac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}$。

2.2 基于 *k*-Means 聚类分析的图像分割

聚类分析的本质就是将具有相同或相似特征的像素点，根据某种相似性度量准则归并为一类，对图像特征进行聚类从而实现图像分割。目前常用的分割方法有 k 均值（k-Means）聚类算法、模糊 c 均值算法、均值漂移算法等。k-Means 聚类算法是最典型的划分聚类法[12]。

k-Means 算法的步骤如下：

1）给定总数据样本集（图像中每一点像素值）$D=\{x_1, x_2, \cdots, x_n\}$，共 n 个像素值。从总数据样本集中任取 k 个样本作为初始的聚类中心 $\{\mu_1, \mu_2, \cdots, \mu_k\}$，即将样本集分为 k 类数据集。

2）计算所有样本 x_i（i=1, 2,⋯, n）与初始聚类中心 μ_j（j=1, 2,⋯, k）的欧氏距离 $D(x_i, \mu_j)=(x_i-\mu_j)^2$；如果满足 $D(x_i, \mu_j)=\min\{D(x_i, \mu_j)\}$，则样本数据 x_i 属于第 j 类数据集 W_j，即 $x_i\in W_j$。

3）对于 j=1, 2, ⋯, k，对 W_j 中所有的样本点重新计算新的聚类中心 $\mu_j=\dfrac{1}{|W_j|}\sum\limits_{x_i\in W_j}x_i$，其中 $|W_j|$ 为第 j 类数据集中样本的个数。

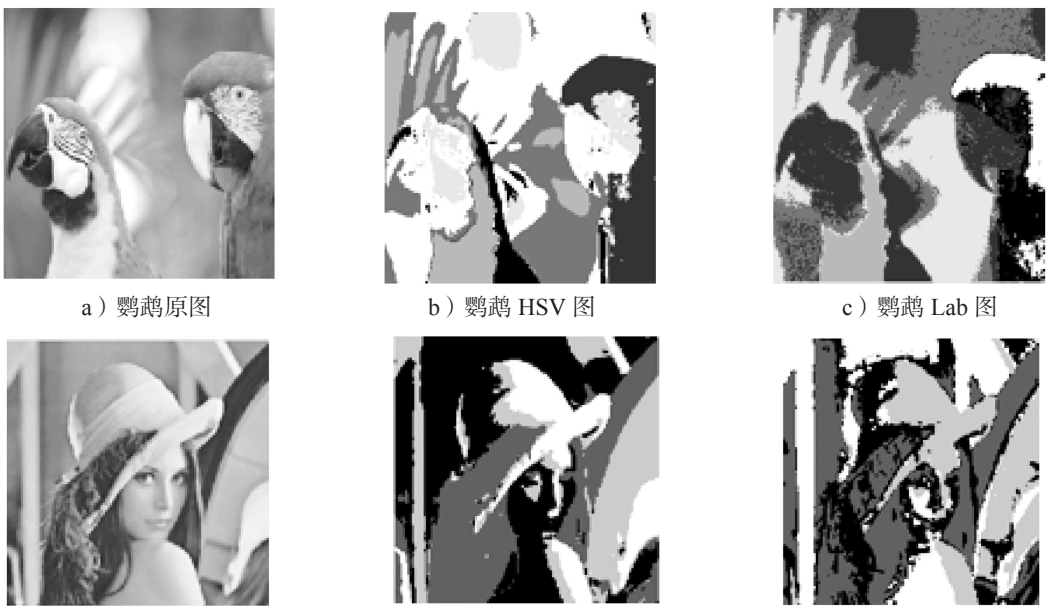
4）如果计算出来的 k 个聚类中心不再发生变化，则得出相应的聚类中心坐标，否则返回第（2）步。

k-Means 聚类算法能实现类内相似度最高和类间相似度最低，且能达到局部最优，因此更适合于彩色图像的分割。首先将采集获得的图像由RGB颜色模型分别转换到Lab和HSV 2个颜色模型进行聚类处理，因为这2个颜色模型的各分量可分离、可单独处理，能够尽量避免光照的影响，更利于进行实验分析。再针对a, b和H, V分别建立二维数据样本，分析实验效果，实验结果如图2和图3所示。

对比图2b和c可以看出，在HSV颜色模型下图像分割更清晰；对比图2e和f可以看出，在HSV颜色模型下图像分割效果明显优于Lab颜色模型。由此可以得出，k-Means 聚类算法在HSV颜色模型下能更好地实现颜色分离。

图3中的sumD是每个类内各点到中心点的距离之和。当聚类中心个数 k=6 时，对比图3a中在HSV颜色模型下与Lab颜色模型下sumD的大小，可以得出，在HSV颜色模型下处理得到的sumD较小，层

次更细腻；当聚类中心个数 k=4 时，对比图 3b 中在 HSV 颜色模型下与 Lab 颜色模型下 sumD 的大小，可以得出在 HSV 颜色模型下处理得到的 sumD 较小，层次也更细腻。由此可以得出，在 HSV 颜色模型下，k-Means 的聚类效果更好，更有利于彩色 QR 码图像的识别。

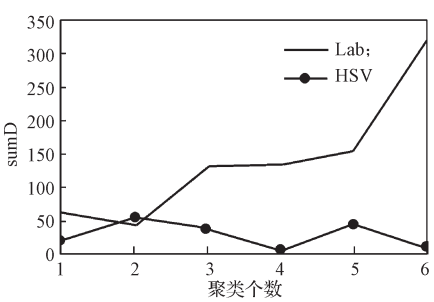


a）鹦鹉原图　b）鹦鹉 HSV 图　c）鹦鹉 Lab 图

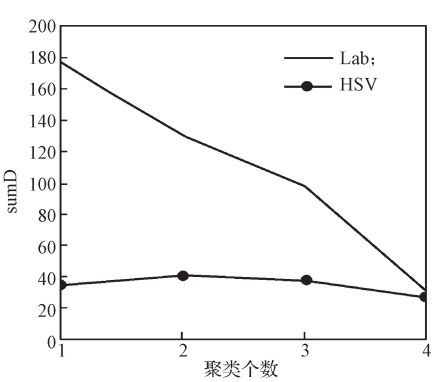
d）lena 原图　e）lena HSV 图　f）lena Lab 图

图 2　不同颜色模型下的聚类效果图

Fig. 2　Clustering effect graph with different color models



a）当 k=6 时鹦鹉图的 sumD



b）当 k=4 时 lena 图的 sumD

图 3　不同颜色模型下的聚类数据结果

Fig. 3　Clustering data results with different color models

2.3　彩色编码模块的配色模型

编码模块颜色的设置直接关系到彩色 QR 码的存储容量及解码精度。从 QR 码存储容量的角度考虑，编码模块的颜色数越多越好；从 QR 码解码精度的角度考虑，编码模块的颜色数越少越好。鉴于彩色 QR 码的应用环境和媒介等诸多因素影响，在 HSV 颜色模型下设置编码模块的颜色时，采用的配色模型为

$$C_i=(V_i, S_i, H_i), \tag{4}$$

式中：C_i 为编码模块第 i(i=0, 1, ⋯, 2^n) 种颜色，其中 n 为彩色 QR 码存储容量扩大的倍数；

V_i 为编码模块第 i 种颜色的明度值，取值范围为 [0, 1]；

S_i 为编码模块第 i 种颜色的饱和度值，取值范围为 [0, 1]；

H_i 为编码模块第 i 种颜色的色调值，取值范围为 [0, 360]。

在 HSV 颜色模型中，色彩的饱和度值 S 对光照强度比较敏感，故在灰色区域设置较少的颜色数，并用明度值 V 按等欧氏距离进行分割。由于饱和度值 S 易受天气、光照等外界环境的不利影响，而色调值 H 较为稳定，对不同光照条件具有不变性，故选取饱和

度值 S_i=1，即纯色。同时，利用稳定的色调值 H 按等欧氏距离来设置颜色，以增强编码模块的辨识度，提高彩色 QR 码的解码精度。基于上述彩色编码模块配色模型，若编码模块设置 2^n 种颜色，则配色规则如下：

当 $0 \leqslant i \leqslant 2^{n-2}-1$ 时，

$$C_i = \begin{cases} V_i = i\dfrac{1}{2^{n-2}-1}, \\ S_i = 0, \\ H_i = 0。 \end{cases} \tag{5}$$

当 $2^{n-2}-1 < i \leqslant 2^n-1$ 时，

$$C_i = \begin{cases} V_i = 1, \\ S_i = 1, \\ H_i = \dfrac{360}{2^n - 2^{n-2}}\left(i - 2^{n-2}\right)。 \end{cases} \tag{6}$$

3 基于光线补偿的 *k*-Means 聚类算法的彩色 QR 码解码

3.1 彩色 QR 码实例

QR 码的存储容量大小取决于每个颜色模块所容纳的比特流的个数。若彩色 QR 码的每个颜色模块表征 k bit，那么此彩色 QR 码需要 2^k 种颜色模块来填充，即码字总数、剩余位、数据码字数、纠错码字数、纠错的块数等变为黑白 QR 码的 k 倍。

彩色 QR 码以黑白 QR 码为基础，保留了黑白 QR 码的功能区、版本信息和格式信息，将数据区域的黑白码块替换为彩色码块。由于功能区、版本信息和格式信息这 3 个区域占 QR 码的面积比例很小，同时为了保证解码的稳定性，所以只需在数据区域提高 QR 码的数据容量。图 4 所示，是基于 *k*-Means 算法的彩色 QR 码解码生成的 16 色彩色 QR 码实例。

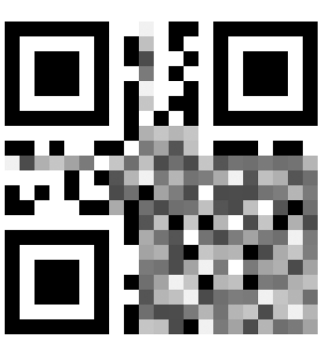


图 4 彩色 QR 码
Fig. 4 Color QR code

3.2 解码实验分析

彩色 QR 码的主要解码过程如下：

1）加载并遍历获取的 png 格式的高容量彩色 QR 码图像，获取 RGB 值。

2）将图像的 RGB 值转换成 HSV 值。

3）进行彩色 QR 码图像预处理。

4）用寻像图形的定位算法检测 3 个位置探测图形，若成功检测到寻像图形，则返回 3 个位置探测图形中心点的坐标值；若有校正图形，则返回校正图形中心点的坐标值。

5）利用寻像图形和校正图形中心点的坐标，进行透视变换，栅格化彩色 QR 码颜色模块。

6）识读颜色映射表，并识读格式和版本信息。

7）运用 RS（reed-solomon）算法进行错误检查，若发现错误则用 RS 算法分块纠错，然后进行步骤 8，若无错则直接进入步骤 8。

8）根据预先设定的颜色映射表，将数据区域内颜色模块与数据信息流一一对应。

9）根据模式指示符和字符计数指示符，进行数据码字译码，恢复存储信息。

10）解码结束。

彩色 QR 码图像的预处理阶段主要实现颜色编码模块的分离。*k*-Means 聚类算法的聚类中心点的个数 k 需要根据图像的特征进行选取，k 值不同会影响算法实现的效率和效果，但由于彩色 QR 码颜色编码模块种类数在实验前是确定的，所以不需要实验选取。

提出基于 HSV 颜色模型的光线补偿的 *k*-Means 聚类算法，与经典的 *k*-Means 聚类算法相比，其主要区别是，该方法首先采用光线补偿法对 HSV 颜色模型的 V 分量进行补偿，然后对图像进行中值滤波处理，在此基础上进行 *k*-Means 聚类处理，使得相似颜色聚集到一个区域，从而解决彩色编码模块的混叠效应。

为验证所提算法的有效性，将其与基于图像直方图法和经典 *k*-Means 聚类算法进行比较，分析它们的所耗时长和解码的正确率。利用上述彩色 QR 码的解码过程，针对如图 4 所示的 16 色彩色 QR 码符号，通过不同的算法处理进行解码，记录所耗时长，并计算解码的正确率，结果如表 2 所示。

由表 2 可知，虽然基于图像直方图法所需时间较短，但由于其对解码的正确率提高得较小，所以此方

法不可取。用手机拍摄采集的彩色 QR 码图像，未采用算法处理时，其译码正确率约为 57.5%~61.4%，这说明光照、天气等环境因素对彩色 QR 码译码的影响较大。采用基于图像直方图法处理的图像，其译码正确率约为 60.1%~65.3%；采用经典 *k*-Means 聚类算法处理的图像，其译码正确率约为 82.7%~83.5%；采用基于 HSV 颜色模型的光线补偿的 *k*-Means 聚类算法处理的图像，其译码正确率有较大提高，约为 86.1%~87.3%。

由此可以看出，采用 HSV 颜色模型中的配色模型进行彩色编码模块的颜色设置，并采用 HSV 颜色模型的光线补偿的 *k*-Means 聚类算法，进行彩色编码模块的色彩分割，聚类个数 *k* 根据颜色编码模块种类来确定，不需要重复计算，识别耗时较短，可以大幅度地降低彩色 QR 码编码模块之间的混叠效应，显著地提高彩色 QR 码译码的正确率。

表 2　不同方法下彩色 QR 码解码正确率

Table 2　Decoding accuracy of color QR codes with different methods

名　称	图像	来　源	算法处理	算法效率 /s	正确率 /%
彩色 QR 码 1		手机拍摄	未处理	0	57.5
			基于图像直方图法	0.36	60.1
			k-Means 算法	0.74	83.5
			基于光线补偿的 *k*-Means 算法	0.81	87.3
彩色 QR 码 2		手机拍摄	未处理	0	61.4
			基于图像直方图法	0.38	65.3
			k-Means 算法	0.84	82.7
			基于光线补偿的 *k*-Means 算法	0.91	86.1

4　结语

综上所述，本文采用基于 HSV 颜色模型的配色模型进行彩色编码模块的颜色设置，降低了编码模块之间的混叠效应，较大程度地消除了彩色 QR 码受天气、光照等外界环境的不利影响。采用基于 HSV 颜色模型的光线补偿的 *k*-Means 聚类算法，进行彩色编码模块的色彩分割，较大地提高了彩色 QR 码解码的正确率。

在信息安全方面，可以通过在 QR 码上嵌入公钥加密和数字签名来保证受众对信息的认证以及确保信息的完整。为了进一步增加 QR 码的存储容量，在今后的研究中可以将专注点放在颜色模块的配色及识别、数据压缩与解压等方面。

参考文献：

[1] 陈　钰 . 面向手机的二维条码识别技术的研究与应用 [D]. 沈阳：东北大学，2005.
CHEN Yu. Research and Application of 2D Barcode Recognition Technology for Mobile Phone[D]. Shengyang：Northeastern University，2005.

[2] 苏　航，傅民仓 . 基于波形分析的彩色 QR 码识别 [J]. 科学技术与工程，2016，16(26)：93−97.
SUN Hang，FU Mincang. Color QR Code Recognition Based on Waveform Analysis[J]. Science Technology and Engineering，2016，16(26)：93−97.

[3] HAN T D，CHEONG C H，LEE N K，et al. Machine Readable Code Image and Method of Encoding and Decoding the Same：United States，7020327[P]. 2002−03−07.

[4] BHARDWAJ N，KUMAR R，VERMA R，et al. Decoding Algorithm for Color QR Code：A Mobile Scanner Application[C]//International Conference on Recent Trends in Information Technology. Chennai：IEEE，2016. DOI：10.1109/ICRTIT.2016.7569561.

[5] KIKUCHI M，FUJIYOSHI M，KIYA H. A New Color QR Code Forward Compatible with the Standard QR Code Decoder[C]//International Symposium on Intelligent Signal Processing and Communications Systems. Naha：IEEE，2014：26−31.

[6] 陈元枝，邓　艳，史绍亮，等 . 基于 Zxing 的彩色 QR 码生成与识别方法 [J]. 桂林电子科技大学学报，2016，36(4)：333−337.
CHEN Yuanzhi，DENG Yan，SHI Shaoliang，et al. Generation and Recognition Method of Color QR Codes Based on Zxing[J]. Journal of Guilin University of Electronic Technology，2016，36(4)：333−337.

[7] 郑河荣，袁远松，赵小敏 . 具有高压缩比汉字编码能力的彩色二维条码及其编码、解码方法：中国，CN200910097065.1[P]. 2009−08−26.

ZHENG Herong，YUAN Yuansong，ZHAO Xiaomin. Color Two-Dimensional Bar Code with High Compression Ratio Chinese Character Coding Ability and Decoding Method Thereof：China，CN200910097065.1[P]. 2009-08-26.

[8] 牛万红，颜惠琴，葛永斌 . 一种彩色 QR 码的设计原理及编解码实现 [J]. 北京理工大学学报，2015，35(10)：1067-1073.
NIU Wanhong，YAN Huiqin，GE Yongbin. A Colored QR Code Design Principle and Encoding-Decoding Implementation[J]. Transactions of Beijing Institute of Technology，2015，35(10)：1067-1073.

[9] [佚　名]. 颜色空间及分类 [EB/OL]. [2013-02-07]. http://blog.csdn.net/zdyueguanyun/article/details/8575865.
[Anon]. Color Space and Classification [EB/OL]. [2013-02-07]. http://blog.csdn.net/zdyueguanyun/article/details/8575865.

[10] [佚　名]. Lab 颜色空间 [EB/OL]. [2016-07-10]. http://www.cvvision.cn/6762.html. [Anon]. Lab Color Space[EB/OL]. [2016-07-10]. http://www. cvvision.cn/6762.html.

[11] 贾　骥 . 道路交通标志检测与识别算法研究 [D]. 北京：北京交通大学，2015.
JIA Ji. Research on Road Traffic Sign Detection and Recognition Algorithm[D]. Beijing：Beijing Jiaotong University，2015.

[12] 陈抒瑢 . 基于特征提取和色彩分割的偏色检测 [D]. 南京：南京大学，2014.
CHEN Shurong. Color Cast Detection Based on Feature Extraction and Color Segmentation[D]. Nanjing：Nanjing University，2014.

Color QR Code Recognition Based on *k*-Means Algorithm

JIA Dan[1]，YOU Fei[1]，ZHANG Qingli[1]，ZENG Zhigao[2]

（1. College of Communication and Art Design，University of Shanghai for Science and Technology，Shanghai 200093，China；
2. School of Computer Science，Hunan University of Technology，Zhuzhou Hunan 412007，China）

Abstract：In order to reduce aliasing effects in color QR decoding and improve the accuracy of color QR code decoding, a *k*-Means clustering algorithm based on HSV color space was proposed. Experiments were conducted to verify the effect of *k*-Means clustering algorithm in three color spaces of RGB, Lab and HSV for choosing the right color space for color QR codes. In HSV color space, the color matching model of color coding module was established according to the principle of equal Euclidean distance, so as to minimize the aliasing effect of color in decoding. In the preprocessing phase of color QR code decoding, *k*-Means clustering algorithm based on ray compensation in HSV color spaces was used to separate color coding module and to improve the decoding accuracy. The result showed that in HSV color space, the *k*-Means clustering effect was the best, with the image region classification effect being the clearest. The color matching model could optimally match the color coding module. Therefore *k*-Means clustering algorithm based on ray compensation could be used in HSV color space to improve the accuracy of color QR decoding. Establishing a reasonable model was to set color encoding modules, and to use *k*-Means clustering algorithm based on ray compensation in HSV color space to achieve color segmentation. It could greatly reduce the aliasing effect between QR color encoding modules and significantly improve the accuracy of color decoding QR.

Keywords：color QR code；*k*-Means clustering algorithm；HSV color model；color matching model of color QR code