

# 分块 K-LT 及查表量化编码的彩色图像压缩方法<sup>①</sup>

牛万红<sup>②\*</sup> 颜惠琴<sup>\*\*</sup> 葛永斌<sup>\*\*</sup>

( \* 宁夏大学远程教育学院 银川 750021 )

( \*\* 宁夏大学数学计算机学院 银川 750021 )

**摘 要** 针对卡洛变换(K-LT)应用于彩色图像压缩时因额外码流开销大而导致编/解码速度慢的缺点,提出了一种最大限度地减少额外码流开销的彩色图像压缩方法:先将 RGB 色彩空间的图像转换到 YCbCr 色彩空间采样,初步去除各颜色分量间的相关性;然后合并各色彩分量,进一步利用分块 K-LT 去除分量内部的信息相关性;最后对 K-LT 的重要系数采取查表的策略进行量化和哈夫曼编码。分块 K-LT 大幅度降低了向量空间的维数及变换矩阵的维数,减少了计算量和额外码流的存储开销;而查表提取 K-LT 重要特征系数的策略进一步节省了码流的存储空间。实验结果表明,该方法不仅编/解码速度快,而且相同压缩比下的峰值信噪比明显高于 JPEG2000 压缩方法。

**关键词** YCbCr 空间, 彩色图像压缩, 分块 K-LT, 查表, 峰值信噪比

## 0 引 言

数字图像的压缩比取决于图像信号中冗余度解除的程度,冗余度可借助相关性度量<sup>[1]</sup>。如果一种变换能将图像信号中像素间的相关性完全解除,从数据压缩来考虑,该变换无疑是一种最佳变换。而卡洛变换(Karhunen-Loeve transform, K-LT)就是这样的一种变换,即建立在统计特性基础上的一种变换,它可将信号在变换域的相关性全部解除,因而是最小均方误差意义下的最佳变换<sup>[2]</sup>。然而,K-LT 应用在图像压缩过程中存在两个不利因素:一是 K-LT 的变换矩阵与输入图像有关,这意味着在图像压缩过程中,除了要把系数矩阵传给解码器,还需要把变换矩阵传给解码器,使得压缩比值大打折扣;二是 K-LT 需要计算协方差矩阵及其特征向量矩阵,计算量相当大而且耗时。因此,K-LT 被研究者们认为只能作为一个衡量其他变换去相关性能的标尺,而很少在图像压缩技术中应用<sup>[3]</sup>。奚宁等<sup>[4]</sup>曾提出把一幅图像划分成不同的组,然后利用 BP 神经网络的方法将各个子图像正确归类,对具有不同统计特性的子图像采用不同的变换矩阵进行 K-LT,以提高整体算法的速度,但算法的空间复杂度较高而且会

增加额外的存储开销。陈嘉臻<sup>[5]</sup>提出了一种基于分段 K-LT 的遥感图像压缩方法,该方法利用分段 K-LT 去除谱间冗余,但空间冗余还需用其它的二维数学变换来消除。本文在以往研究的分块 K-LT 算法<sup>[6]</sup>基础上,针对 K-LT 应用在图像压缩过程中的缺陷,提出了一种 YCbCr 空间下,基于分块 K-LT 及查表量化、编码的彩色图像压缩方法,以最大限度地减少额外码流的开销,实现图像压缩。实验表明,该方法相比 JPEG2000 压缩方法<sup>[7-9]</sup>具有更好的压缩性能。

## 1 算法理论基础

### 1.1 色彩空间转换

彩色图像在 RGB 色彩空间中的 3 个分量之间存在着高度的相关性,即色彩冗余信息。图像压缩的目的就是要尽可能地消除图像中的全部冗余信息,最大限度地降低表示图像信息所需的数据量。去除彩色图像 3 个分量之间的冗余信息,一般的方法是通过色彩空间变换来实现的。而把 RGB 色彩空间转换到 YCbCr 色彩空间是图像处理技术中经常用到的方法,其中 Y 代表亮度信息分量,Cb、Cr

① 国家自然科学基金(11361045),宁夏自然科学基金(NZ13050)和宁夏高等学校科学技术研究(413-0082)资助项目。

② 男,1966 年生,硕士,副教授;研究方向:图形处理与多媒体技术;联系人,E-mail: niu\_ wh@nxu. edu. cn  
(收稿日期:2014-02-24)

分别代表蓝色色差分量和红色色差分量。RGB 转换为 YCbCr 所用的变换公式为<sup>[10]</sup>

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.578 & 0.114 \\ 0.500 & -0.4187 & -0.0813 \\ -0.1687 & -0.3313 & 0.500 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} \quad (1)$$

由于人的视觉对亮度分量的敏感程度比对色度分量大,在 YCbCr 色彩空间下,通过对  $Cb$ 、 $Cr$  色度分量进行采样,可降低各色彩分量之间的冗余信息,同时可节省图像编码所需要的存储开销。采样格式选取流行的 4:2:0 格式<sup>[11]</sup>,在这种格式下  $Cb$  和  $Cr$  矩阵在水平和垂直方向上的分辨率是  $Y$  矩阵的一半,如图 1 所示。

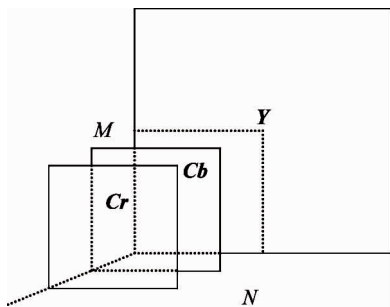


图 1 YCbCr 空间 4:2:0 采样

将 RGB 色彩空间转换成 YCbCr 色彩空间进行采样,仅仅是一个初步去除各色彩分量间相关性的过程,采样后各分量之间的相关性有所削弱,但各分量内部仍留有很大的视觉相似性<sup>[12]</sup>,可进一步利用分块 K-LT 去除各分量内部的信息相关性。

## 1.2 分块 K-LT 原理

假设图像矩阵为  $D$ ,对矩阵  $D$  进行分块 K-LT。子块的分辨率定为  $m \times n$ ,为了降低特征空间的维数,减少协方差矩阵及其特征值和特征向量矩阵的计算量,提高 K-LT 的运算速度, $m$  和  $n$  的取值要小些,分割的子图像块数  $Q$  需满足  $Q \gg (m \times n)$ 。然后将每个子图像的像素向量化,即采用行堆叠或列堆叠的方法分别形成一个  $m \times n$  维向量<sup>[13]</sup>,则  $Q$  个子图像形成的所有向量构成向量组  $x$ ,即  $x: x_1, x_2, \dots, x_Q$ ,矩阵形式表示为

$$X = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{Q1} \\ x_{12} & x_{22} & \cdots & x_{Q2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1mn} & x_{2mn} & \cdots & x_{Qmn} \end{bmatrix} \quad (2)$$

向量组  $x$  的均值向量为

$$\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_{mn})^T \quad (3)$$

其中  $\bar{x}_j = \frac{1}{Q} \sum_{i=1}^Q x_{ij}$ ,  $j = 1, 2, \dots, m \times n$ 。其协方差矩阵  $S_x$  可由下式近似计算:

$$S_x = \frac{1}{Q-1} \sum_{i=1}^Q (x_i - \bar{x})(x_i - \bar{x})^T \quad (4)$$

为了获得  $S_x$  的无偏估计,这里  $Q-1$  代替  $Q$ 。可知,  $\bar{x}$  是  $m \times n$  维的向量,  $S_x$  是  $mn \times mn$  的方阵。

求得协方差矩阵后,可求出  $m \times n$  个特征值,并对特征值按大小进行排序,如  $\lambda_1 > \lambda_2 > \dots > \lambda_{mn}$ 。再进一步求出特征值  $\lambda_j$  对应的归一化特征向量  $e_j$  构成的特征向量矩阵,即变换矩阵

$$A = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1mn} \\ e_{21} & e_{22} & \cdots & e_{2mn} \\ \vdots & \vdots & \ddots & \vdots \\ e_{mn1} & e_{mn2} & \cdots & e_{mnmn} \end{bmatrix} \quad (5)$$

变换矩阵  $A$  的阶数也就相应地为  $m \times n$ 。

K-LT 表达式为

$$y = A(x - \bar{x}) \quad (6)$$

逆变换为

$$x = A^T y + \bar{x} \quad (7)$$

从上述原理可知,分块 K-LT 并非是对每一个子块单独进行 K-LT,而是将每一个子块像素向量化后重组为一个新的 2 维矩阵再进行 K-LT。其中子块的分辨率决定着变换矩阵特征空间的维数大小。子块越小,则特征空间的维数越低,算法的执行时间也越短,去相关效率也会降低。反之,子块越大,则特征空间的维数越大,算法的执行时间也越长,虽有极好的去相关效率,但图像编码时作为附加数据量的变换矩阵的尺寸也较大,不利于压缩比的提高。因此,子块的分辨率由变换矩阵的尺寸而确定,通常将变换矩阵的数据存储空间预先控制为原图像的 1% 以内,则不会对编码的压缩比产生太大的影响。

如此设计的分块 K-LT 算法作用于彩色图像压缩时,相比文献[6]提出的大分块 K-LT 算法,有着更进一步的优势。原因是:本文设计的分块 K-LT 算法去除的是每一个子图像内部像素之间的相关性,而文献[6]提出的大分块 K-LT 算法去除的是子图像之间对应像素的相关性。通常情况下,相邻像素间的信息相关性要高于间隔距离较大的像素间的信息相关性。因此,本文设计的分块 K-LT 算法的去相关效率略高于文献[6]提出的大分块 K-LT 算法的去相关效率。

2 编/解码原理

YCbCr 色彩空间下分块 K-LT 的彩色图像编码原理框图如图 2 所示。

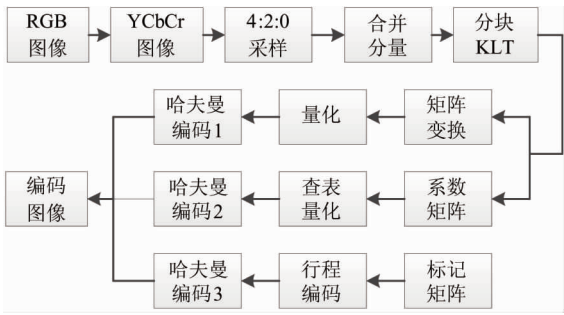


图 2 编码原理框图

编码系统中,如果对 YCbCr 色彩空间下的 3 个色彩分量分别进行分块 K-LT,则产生 3 个变换核矩阵和 3 个系数矩阵,再分别进行量化、编码。实验结果显示,相同压缩比下,峰值信噪比 (peak signal to noise ratio, PSNR) 基本持平或略低于 JPEG2000 方法。为了节省码流的存储开销,文中采取的策略是将 3 个分量重新合并为一个 2 维矩阵,共同参与分块 K-LT,如图 3 所示。这样 3 个分量能够共同使用 1 个变换核矩阵并且仅产生 1 个变换系数矩阵,极大地降低了码流的存储开销。由分块 K-LT 原理可知,算法去除的是各子图像内部像素间的相关性,所以这种做法并不影响各色彩分量内部相关性的去除。

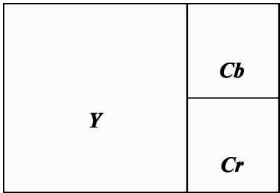


图 3 色彩分量合并

K-LT 本身不能起到直接压缩图像的目的,它只是起到一个去除图像信息相关性的作用。变换后的重要特征系数主要集中在前面  $P$  行,但仍有相当一部分非重要信息掺杂其中,占了较大的存储空间。如果简单地采取截断的方法进行量化和编码,不利于提高压缩比,因此本文在量化之前预先设定一个阈值  $T$ ,采取查表抽取重要特征系数的策略<sup>[14]</sup>来节省存储空间。阈值的大小决定着压缩比的大小,实验表明,舍去  $T \leq 0.2$  的系数,对图像恢复几乎不产生影响,可实现近似无损压缩;反之,随着  $T$  值的增

大,图像的压缩比也越来越大。主要算法步骤描述如下:

(1) 设系数矩阵  $F$  的尺寸为  $M_1 \times N_1$ ,预先设定一个阈值  $T$ ,用来界定  $F$  中待量化的系数,也用来调节压缩比的大小;再预先分配一个尺寸为  $M_1 \times N_1$  的零矩阵  $Z$ ,用来标记  $F$  中被选出的系数。

(2) 查找  $F$  中绝对值大于或等于阈值  $T$  的系数存入矩阵  $E$  中,同时修改零矩阵  $Z$  中相应位置的元素值为 1,即:

$E = F(\text{find}(\text{abs}(F) \geq T))$ ; //选出待量化的系数;

$Z(\text{find}(\text{abs}(F) \geq T)) = 1$ ; //构造标记矩阵。

直观效果如图 4 所示。

(3) 对矩阵  $E$  中的系数采取均匀量化及哈夫曼 (Haffman) 编码。

(4) 根据标记矩阵  $Z$  中元素为 0 或 1 的特点,先对其进行行程编码 (RLE),再进行哈夫曼编码,达到最佳编码效果。

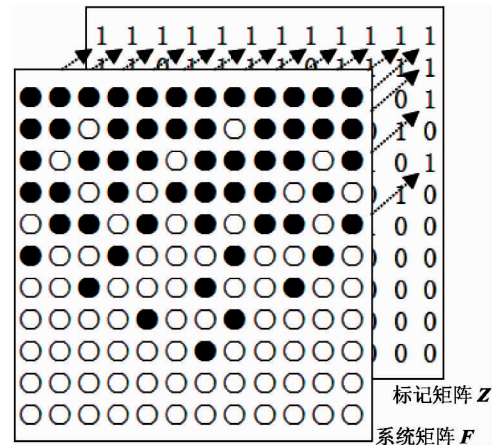


图 4 构造标记矩阵

图 4 中“●”表示绝对值大于阈值  $T$  的系数,对应于矩阵  $Z$  中的元素 1;“○”表示绝对值小于阈值  $T$  的系数,对应于矩阵  $Z$  中的元素 0。

解码和编码过程是完全对称的,只需要将输出改成输入即可。

3 实验结果

选用 3 种具有不同纹理特征的彩色位图图像 (分辨率为  $512 \times 512$ ) 作为压缩对象,在 Core2 Duo T7100 微机上对图像的编/解码实验,利用压缩比和峰值信噪比 (PSNR) 2 项指标来衡量算法的有效

性。在相同压缩比的情况下,重构图像的 PSNR 越大,表明算法的性能越好。假设  $f(i, j)$  是 RGB 图像的一个色彩分量,  $f'(i, j)$  是解码后相应的色彩分量,  $M$  和  $N$  分别为色彩分量的行数和列数,即每个色彩分量的分辨率为  $M \times N$ ,则峰值信噪比定义为

$$PSNR = 10 \times \log\left(\frac{255^2}{MSE}\right) \tag{8}$$

MSE 为均方误差,表达式为

$$MSE = \frac{1}{M \times N \times 3} \sum_{f \in \{R, G, B\}} \sum_{i=1}^M \sum_{j=1}^N (f'(i, j) - f(i, j))^2 \tag{9}$$

表 1 分块 K-LT 和 JPEG2000 的 PSNR 比较

| 压缩比 Cr | 图 shan512. bmp |           | 图 lena512. bmp |           | 图 hua512. bmp |           |
|--------|----------------|-----------|----------------|-----------|---------------|-----------|
|        | 分块 K-LT        | JPEG-2000 | 分块 K-LT        | JPEG-2000 | 分块 K-LT       | JPEG-2000 |
| 22.3   | 53.45          | 53.44     | 52.28          | 51.96     | 52.53         | 52.43     |
| 40.1   | 52.64          | 49.48     | 46.41          | 45.65     | 50.38         | 48.59     |
| 60.0   | 49.99          | 45.26     | 43.59          | 41.86     | 48.63         | 46.33     |
| 80.2   | 45.29          | 41.54     | 42.41          | 39.75     | 46.39         | 44.72     |
| 100.1  | 42.69          | 38.62     | 41.50          | 38.44     | 44.16         | 43.11     |
| 110.0  | 41.07          | 37.55     | 40.97          | 37.93     | 43.18         | 42.48     |

表 1 给出了 3 幅图像分别利用分块 K-LT 压缩方法和 JPEG2000 压缩方法进行 6 种压缩比实验的峰值信噪比比较结果。可以看出,在压缩比较低或近无损压缩的情况下(如  $Cr = 22.3:1$ ),分块 K-LT 压缩方法和 JPEG2000 压缩方法的峰值信噪比基本持平;随着压缩比的增大或有损压缩的情况下,前者的峰值信噪比明显高于后者,尤其是对于图像 shan512. bmp,因其纹理较为稀疏,最大差值可达到 4.73 dB( $Cr = 60:1$ )。说明在相同峰值信噪比的情况下,分块 K-LT 压缩方法有着更大的压缩比。

部分主观视觉效果对比( $Cr = 110:1$ )如图 5、图 6和图 7 所示。

从图 5、6、7 可以看出,3 种不同纹理特征的图像,在压缩比等于 110:1 时,两种压缩方法都有着良好的重构视觉效果;当压缩比大于 110:1 时,纹理特征较为稀疏的图像 shan512. bmp,利用分块 K-LT 压缩方法的重构图像会逐渐产生细微的块效应,主要是由于该方法在量化前舍去了大量的非重要特征系数造成的。如果在舍去非重要特征系数之前采取非均匀量化,则视觉效果要好些,但涉及到多个量化阈值的选取,过程难以调控,本文未采用。而纹理特征表现丰富的图像 lena512. bmp 和 hua512. bmp,在压缩比为140:1的情况下,两种压缩方法重构图像的

其中  $f \in \{R, G, B\}^{[15]}$ 。  
首先将所选图像转换到 YCbCr 色彩空间并按 4:2:0 格式采样,再合并各色彩分量并采取  $8 \times 8$  的分块。由分块 K-LT 原理可知,变换矩阵的尺寸为  $64 \times 64$ ,其存储空间是原图像存储空间  $512 \times 512 \times 3$  的 0.52%,这对减少额外码流的开销极为有利。JPEG2000 压缩方法同样是在 YCbCr 色彩空间下按 4:2:0 格式采样,所不同的是要对 3 个色彩分量分别进行小波变换,再采取相应的措施进行量化、编/解码。实验结果见表 1,  $Cr$  表示压缩比,峰值信噪比(PSNR)的计算是在 RGB 空间进行的。



图 5 shan512. bmp 视觉效果对比



图 6 lena512. bmp 视觉效果对比

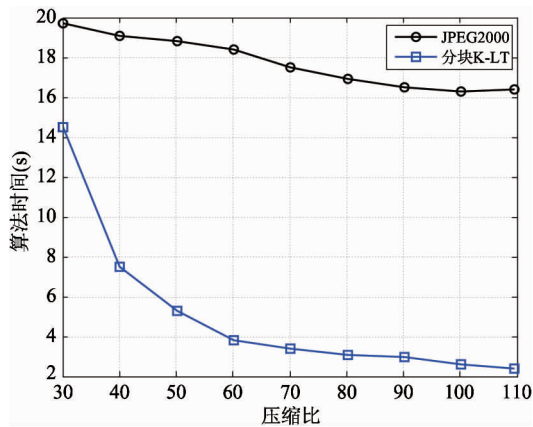


图 7 hua512. bmp 视觉效果对比

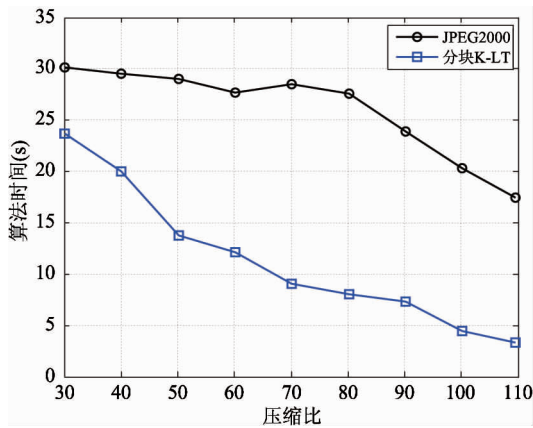


视觉效果仍然能保持良好,压缩比大于 150:1 后,均开始失真。

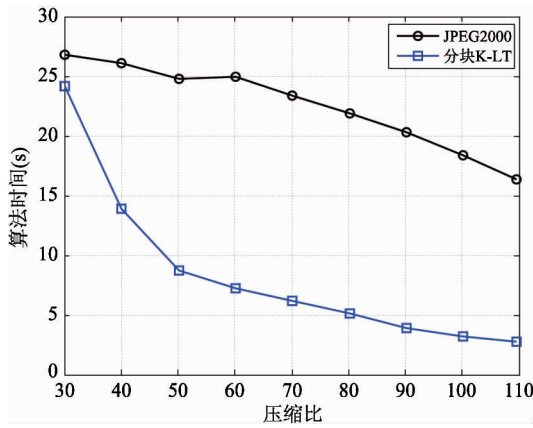
为了进一步验证分块 K-LT 压缩方法的有效性,再从编/解码速度方面来考察该方法的性能。图 8(a)、(b)、(c) 分别绘制出了上述 3 种图像利用该方法或 JPEG2000 压缩方法进行编/解码在不同压缩比下各自所用时间的对比曲线。



(a) shan512. bmp 图像



(b) lena512. bmp 图像



(c) hua512. bmp 图像

图 8 分块 K-LT 与 JPEG2000 的编/解码速度比较

由于不同纹理特征的图像,在同一压缩比下编/解码时间有所差异,所以纹理特征丰富的图像一般用时较多,反之较少。可以看出 3 幅图像分别利用这 2 种方法进行编/解码,其时间都随着压缩比的增大而减少,当压缩比大于 50:1 或 60:1 时,分块 K-LT 方法的编/解码时间基本在 3 ~ 12s 以内完成,几乎是 JPEG2000 方法编/解码时间的 1/6 ~ 1/3。这充分说明,有损压缩的情况下,本文提出的分块 K-LT 结合查表抽取 K-LT 重要特征系数的方法,占用存储空间较小,有着更快的编/解码速度。

另外,文献[6]提出的大分块 K-LT 算法,利用上述素材进行图像压缩时,如果不结合查表量化编码的策略,则压缩性能明显不如本文提出的方法,限于篇幅不再给出详细的数据比较。

## 4 结 论

传统的卡洛变换(K-LT)应用于彩色图像压缩时,由于变换矩阵较大导致额外的码流开销也较大,不但编/解码速度慢而且压缩比也较低。本文提出的 YCbCr 空间下,基于分块 K-LT 及查表量化、编码的彩色图像压缩方法,有效地解决了这一问题,其优势在于:(1)降低了特征空间的维数,减少了计算量,大幅度提升了变换速度;(2)仅使用了一个低阶的变换矩阵,减少了额外码流的存储开销,大大提高了图像的压缩比;(3)采取查表提取 K-LT 系数矩阵中重要特征系数的策略进行量化和编码,节省了码流的存储空间,进一步提高了图像的压缩比;(4)充分利用 K-LT 的统计特性优势及其最小均方误差意义下最佳变换的优点,在有损压缩的情况下进行彩色图像编码。本文方法相比 JPEG2000 方法,相同压缩比下具有更高的峰值信噪比。

## 参考文献

[ 1 ] Kervrann C, Boulanger J, Pecot T, et al. Discriminant random field and patch-based redundancy analysis for image change detection. In: Proceedings of the International Conference on Machine Learning for Signal Processing, Grenoble, France, 2009. 1-6

[ 2 ] Kountchev R, Kountcheva R. Image compression based on the Karhunen-Loeve color transform with Inverse Pyramid decomposition. In: Proceedings of the 10th International Conference on Telecommunication in Modern Satellite Cable and Broadcasting Services, Nis, Serbia, 2011. 315-324

[ 3 ] Ragab A S, Mohamed A S A, Hamid M S. Efficiency of

- analytical transforms for image compression. In: Proceedings of the 15th International Conference on Radio Science Conference, Cairo, Egypt, 1998. 1-10
- [ 4 ] 奚宁, 翁默颖. 使用 K-L 变换进行图像压缩的一种新方法. 华东师范大学学报(自然科学版), 1996, 3(3): 63-71
- [ 5 ] 陈嘉臻, 闫敬文. 基于分段 K-L 变换的遥感图像压缩. 福建电脑, 2007, 05: 17 + 14
- [ 6 ] 牛万红, 赵静, 葛永斌. 基于 K-LT 的高分辨率图像的分块算法. 济南大学学报(自然科学版), 2014, 03: 198-203
- [ 7 ] Skodras A, Christopoulos C, Ebrahimi T. The JPEG-2000 still image compression standard. *Signal Processing Magazine*, 2001, 18(5): 36-58
- [ 8 ] Khan E, Ghanbari M. Wavelet based efficient color image coding technique. In: Proceedings of the International Conference on Image Processing, New York, USA, 2002. 1241-1244
- [ 9 ] Rabbani M, Joshi R. An overview of the JPEG2000 still image compression standard. *Signal Processing: Image Communication*, 2002, 17(1): 45-48
- [ 10 ] 刘富强, 王新红, 宋春林. 数字视频图像处理与通信. 北京: 机械工业出版社, 2010. 17
- [ 11 ] Imany P, Yazdanpanah M, Miran S, et al. A devised approach to optimize color space transformation for image compression. In: Proceedings of the 20th International Conference on Telecommunications Forum, Belgrade, Serbia, 2012. 1737-1740
- [ 12 ] Boke S B, Diri B, Yavuz Z K. Lossy compression of color images based on color spaces. In: Proceedings of the 18th International Conference on Signal Processing and Communications Applications Conference, Diyarbakir, Turkey, 2010. 609-612
- [ 13 ] Chuohao Y, Yih H T, Zhengguo L, et al. Mode-dependent fast separable KLT for block-based intra coding. In: Proceedings of the International Symposium on Circuits and Systems, Rio de Janeiro, Brazil, 2011. 621-624
- [ 14 ] Ngoc-Mai N, Xuan-Tu T, Vivet P, et al. An efficient Context Adaptive Variable Length coding architecture for H.264/AVC video encoders. In: Proceedings of the International Conference on Advanced Technologies for Communications, Hanoi, Vietnam, 2010. 158-164
- [ 15 ] Yoo J C, Ahn C W. Image matching using peak signal-to-noise ratio-based occlusion detection. *Image Processing*, 2012, 6(5): 483-495

## A color image compression method based on block K-LT and look-up-table quantization coding

Niu Wanhong<sup>\*</sup>, Yan Huiqin<sup>\*\*</sup>, Ge Yongbin<sup>\*\*</sup>

(<sup>\*</sup> School of Remote Education, Ningxia University, Yinchuan 750021)

(<sup>\*\*</sup> School of Mathematics and Computer, Ningxia University, Yinchuan 750021)

### Abstract

Aiming at the problem that when the Karhunen-Loeve transform (K-LT) is applied to color image compression, the overhead of additional code stream brings about a slow encoding/decoding speed, a color image compression method for reducing the overhead of additional code stream to a minimum is presented. The method takes the steps below: First, an image in the RGB color space is transformed into the YCbCr color space for sampling and preliminary removal of the correlation between each color component. Then, these color components are again merged so as to further make use of the block K-LT to remove the internal information relevance. Finally, the important characteristic coefficients of the K-LT are quantified and encoded in the Huffman way by adopting the look-up-table strategy. The block K-LT significantly reduces the dimension of the vector space and the transformation matrix, so it reduces the amount of calculation and the store-overhead of additional code stream; on the other hand, the look-up-table strategy saves further on the store-space of code stream. The experimental results show, in the case of the same compression ratio, that the method's encoding / decoding speed is not only faster, but also the peak signal to noise ratio is significantly higher than that of the JPEG2000 compression method.

**Key words:** YCbCr space, color image compression, block karhunen-loeve transforms, look-up-table, peak signal to noise ratio