

一种具有旋转不变性的压缩域纹理图像分类方法¹

黄祥林 沈兰荪

(北京工业大学信号与信息处理研究室 北京 100022)

摘 要 该文提出了一种具有旋转不变性的、基于 DCT 压缩域的纹理图像分类方法。此方法主要是在 DCT 压缩域中合理组织频率系数,使其具有与小波分解相类似的图像多分辨率结构,由此设计了抗旋转能力强的纹理图像分类算法。这些操作直接在游程长编码(RLE)形成的码流上进行,处理的数据量较少,处理速度快。实验结果表明,这种分类方法具有较高的准确性。

关键词 压缩域图像处理, DCT 变换, 纹理图像分类, 旋转不变性

中图分类号 TN911.73

1 引 言

在纹理图像分类算法^[1-5]中,基于多分辨率的小波方法具有明显的优越性^[5]。但是目前大多数的图像数据是以压缩格式进行存储和传输的,对已压缩的纹理图像分类时,必须先对压缩图像解码,然后才能进行分类处理。面对大量的压缩图像,这种解压操作要占用很多时间,影响了整个系统的效率。因此,如何在不解码或尽量少解码的情况下直接进行基于压缩域的纹理图像分类就成了一个研究方向^[6-8]。

现有大多数图像压缩标准都采用了基于 DCT 变换的编码算法(如 JPEG, H.261/263, MPEG-1/2 等),因此基于 DCT 压缩域的纹理图像分类算法受到了广泛的研究。文献[4,5]把 DCT 的每个通道能量作为纹理分类的特征(对于 8×8 DCT,共有 64 个特征)进行分类,由于每个通道的变化都影响分类结果,因此其抗噪能力比较差。文献[8]仔细分析了 DCT 系数的区域分布特点,将整个 DCT 系数组合成能反映纹理方向性(水平、对角、垂直等方向)的几个区域,使用这种具有区域方向性的能量特征进行纹理分类,其试验结果表明有很好的抗噪能力。但这两种方法都不具有抗旋转的能力,因为他们选取的分类特征是具有方向性的。本文在文献[8]的基础上,进一步探索 DCT 系数区域所具有的特性,提出了一种具有旋转不变性的压缩域纹理图像分类方法,主要利用多分辨率技术在纹理分类中的优点,首先将 DCT 系数组合成具有与三级小波分解相似的多分辨率结构,然后,提取各分辨率下所有方向分量的综合特征进行纹理分类。由于分类特征是对各个方向分量的综合,因此本算法具有一定的抗旋转能力,试验结果也证明了这一点。下面分别对 DCT 域的多分辨率特性、具有抗旋转性的纹理分类方法等分别进行论述,最后给出试验结果和结论。

2 DCT 域的多分辨率特性

JPEG, H.261/263, MPEG-1/2 等图像压缩标准都采用了 8×8 的 DCT 变换,其基本过程如图 1。首先把原始图像分成若干 8×8 的像素块,对每个像素块进行 DCT 变换,然后进行量化、Zigzag 排序、游程长编码(Run Length Encoding, RLE)、熵编码等操作,最终形成压缩码流。

为了构造 DCT 变换域的多分辨率特性,这里参考一下对 8×8 图像进行三级小波分解的情况,分解后得出了 10 个子带: LL3, LH3, HL3, HH3 等 4 个子带均只有 1 个系数,

¹ 2001-03-30 收到, 2001-10-18 定稿

国家自然科学基金(60172045)、863 计划(2001AA121061)、北京市自然科学基金(4002002)资助

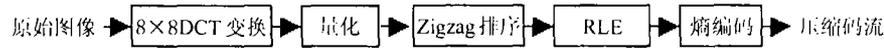


图 1 基于 DCT 变换的压缩编码算法的基本编码框图

LH2, HL2, HH2 等 3 个子带均含 4 个系数, 而 LH1, HL1, HH1 等 3 个子带均含 16 个系数。根据这种方式, 对 DCT 变换的系数也进行 10 个区域的划分: B4, B3-1, B3-2, B3-3, B2-1, B2-2, B2-3, B1-1, B1-2, B1-3, 其所含系数的个数分别为 1, 4, 16, 与小波的各个子带相对应, 如图 2 所示。

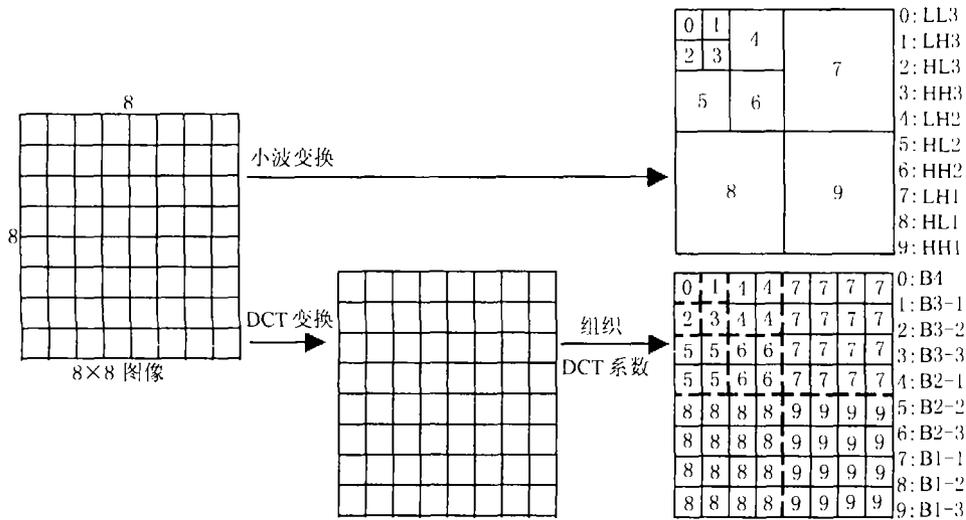


图 2 8 × 8 图像的小波变换与 DCT 变换

对于整幅图像, 把所有的 8 × 8 DCT 块中相对应的区域组织在一起 (如图 3), 即所有的 B4 组成区域 A4, 所有的 B3-1 组成区域 A3-1, ..., 所有的 B1-3 组成区域 A1-3。实际上, 区域 A4 就是整个图像的低频近似 (即所有 DCT 直流系数 B4 所组成的图), 正好与三级小波的 LL3 相对应; 而区域 A3-1, A2-1, A1-1 主要代表了原图像中不同频率段上水平方向的频率变化 (垂直方向的边缘、轮廓、纹理等), 与三级小波 LH3, LH2, LH1 相对应; 区域 A3-2, A2-2, A1-2 则主要集中了原图像中各频率段上垂直方向频率变化 (水平方向的边缘、轮廓、纹理等), 与三级小波的 HL3, HL2, HL1 相似; 区域 A3-3, A2-3, A1-3 基本上原图像中各频率段上对角方向的频率变化 (对角方向的边缘、轮廓、纹理等), 与三级小波的 HH3, HH2, HH1 一致。因此, 通过这种方法对 DCT 系数进行重新组织, 就形成了与三级小波分解基本相似的具有方向性的多分辨率结构的形式。为了更直观的比较, 这里选用一幅轮廓分明的人工图像如图 4(a), 其三级小波分解图像如图 4(b), 经 8 × 8 DCT 变换后重新组织系数得到的多分辨率形式如图 4(c)。从图可知, 它们之间确实具有很强的对应关系。

DCT 域的这种与小波分解相似的具有方向性的多分辨率特性增大了 DCT 变换所具有的处理能力, 非常有利于直接在压缩域进行图像处理, 对于那些采用 DCT 压缩 (如 JPEG, MPEG-1/2 等) 的图像数据, 可以在尽量少量解码的情况下, 进行诸如纹理分类、图像检索、轮廓提取等处理。

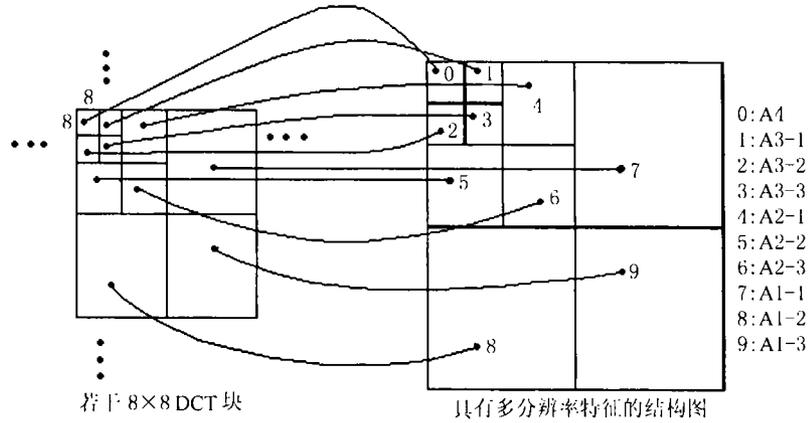


图 3 整幅图像中 DCT 块系数区域的组织

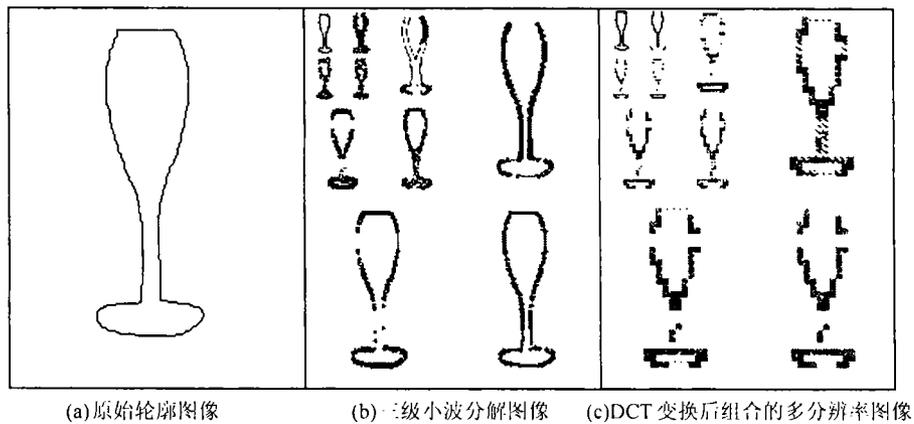


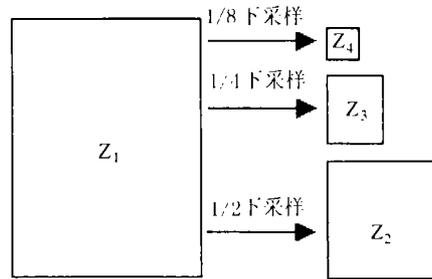
图 4 DCT 变换与小波分解的多分辨率结构图比较

3 具有抗旋转性的纹理分类方法

文献 [2, 3] 采用小波进行纹理分类时, 考虑了各个分辨率的方向性, 文献 [4, 5, 8] 在 DCT 压缩进行分类时, 强调的也是方向性, 这些分类方法都缺乏抗旋转的能力, 实际应用时要受到一定的限制。这里想通过 DCT 所具有的多分辨率特性, 综合各方向性的特征, 使分类算法具有一定的抗旋转能力。

在 DCT 域的多分辨率结构图中, 把同一分辨率的不同方向区域组合在一起, 即把 A3-1, A3-2, A3-3 组合为 A3; 把 A2-1, A2-2, A2-3 组合为 A2; 把 A1-1, A1-2, A1-3 组合为 A1。这样, 新组合的区域 A3, A2, A1 只具有多分辨率结构, 不具备方向性特征。在以 A4, A3, A2, A1 的区域特征进行纹理分类时, 就消除了对方向的敏感性。

在具体的算法设计时, 为了更有效地提取图像的纹理特征, 我们首先以 8×8 DCT 块为单位, 对整个图像的 DCT 域进行 3 次下采样, 即分别进行 $1/2, 1/4, 1/8$ 下采样得到 Z_2, Z_3, Z_4 , 这样在压缩域就形成了具有 4 种分辨率的结构 Z_1, Z_2, Z_3, Z_4 , 如图 5 所示。

图5 以 8×8 DCT 块为单位的 DCT 域 4 种分辨率结构

对每个 Z_i , 分别进行 DCT 域的多分辨率组合, 最后得到 $A_4^{(i)}, A_3^{(i)}, A_2^{(i)}, A_1^{(i)} (i = 1, 2, 3, 4)$ 共 16 个区域。

在纹理分类算法中, 一般采用均值或 / 和方差作为分类特征^[3-5]。在试验中我们发现使用方差比较合适, 它较好地反映了纹理图像的粗糙度等特征。用 F 表示每个区域的方差, 下面以 Z_1 为例, 求其 4 个区域的方差 $F_4^{(1)}, F_3^{(1)}, F_2^{(1)}, F_1^{(1)}$:

设 Z_1 共有 N 个 DCT 块, P_{ij}^m 表示第 m 块中第 i 行、第 j 列的 DCT 系数, 则

$$F_4^{(1)} = \frac{1}{N} \sum_{m=1}^N (P_{00}^m - \bar{P}_0)^2 \quad (1)$$

$$F_k^{(1)} = \frac{1}{N \cdot G_k} \sum_{m=1}^N \sum_{A_k} (P_{ij}^m - \bar{P}_k)^2, \quad k = 1, 2, 3 \quad (2)$$

其中 $\bar{P}_0 = \frac{1}{N} \sum_{m=1}^N P_{00}^m$, $\bar{P}_k = \frac{1}{N \cdot G_k} \sum_{m=1}^N \sum_{A_k} P_{ij}^m$, $k = 1, 2, 3$, G_k 为 A_k 中的系数个数, 其值为 48、12、3 ($k=1, 2, 3$)。

综合 Z_1 、 Z_2 、 Z_3 、 Z_4 中各区域的方差值, 最后形成一个 16 维的特征矢量 $V = \{F_0^{(1)}, F_1^{(1)}, F_2^{(1)}, F_3^{(1)}, F_0^{(2)}, F_1^{(2)}, F_2^{(2)}, F_3^{(2)}, F_0^{(3)}, F_1^{(3)}, F_2^{(3)}, F_3^{(3)}, F_0^{(4)}, F_1^{(4)}, F_2^{(4)}, F_3^{(4)}\}$ 。对于两幅具有 DCT 压缩格式的纹理图像 p, q , 用欧氏距离计算其相似性:

$$d(p, q) = \sqrt{\sum_{i=1}^{16} (V_i^p - V_i^q)^2} \quad (3)$$

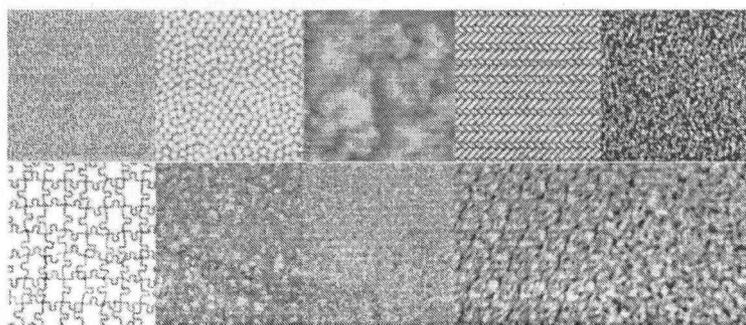
这样, 根据纹理图像之间的相似性就可以进行纹理图像的分类了。

由于分类算法是直接压缩域进行的, 还必须考虑一些特殊处理^[8]。在图 1 中, 熵编码所形成的压缩码流并非字节对齐, 不利于计算机处理, 因此基于 DCT 压缩域的处理通常先进行熵解码再对 RLE 码流进行处理^[6]。处理时可直接跳过 RLE 码流中的 run length 个零值, 只对非零值进行运算, 因此运算量比较少, 节约了处理时间。具体的处理方法与文献 [8] 中 3.4 节的论述相似, 这里不再赘述。

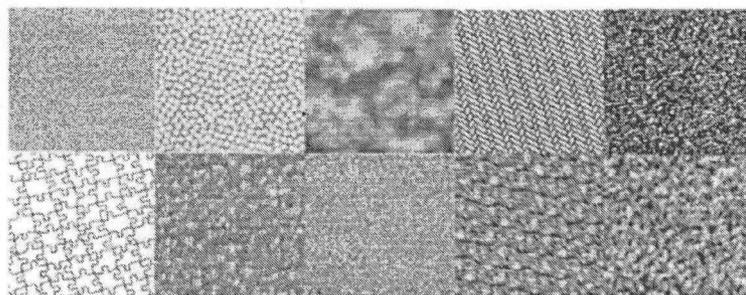
4 实验结果与结论

根据上面的讨论, 作者对 JPEG/DCT 压缩的纹理图像进行了实验模拟。

纹理图像取自 Adobe 纹理图像库, 共十类: Bark, Caviar, Cloud, Denimz, Frznrain, Puzzle, Rustflks, Strands2, Thckhair, Weave3, 分辨率为 512×512 。为了获得足够的训练样本和测试样本, 将原始的 512×512 图像以 $1/2$ 抽样形成 4 幅 256×256 的图像共 40 幅。对所有的图像进行 JPEG 压缩 (DCT 量化矩阵、质量因子一定, SNR 的平均值为 28.12dB, 恢复图像如图 6(a) 所示), 以便在压缩域直接进行纹理分类, 将其中的 20 幅作为训练样本 (这里采用最简单的均值方法求每类的类中心矢量, 即每类 2 幅训练图像特征矢量的均值), 剩下的 20 幅作为测试样本。为了验证此分类方法的抗旋转能力, 利用 Adobe Paint-Shop 中的图像旋转方法, 给所有的 40 幅图像以 $15^\circ, 30^\circ, 45^\circ, 60^\circ, 75^\circ, 90^\circ, 105^\circ, \dots$ 等每隔 15° 的间隔旋转图像共 $10 \times 4 \times 24 = 960$ 幅, 每类 96 幅图像 (图 6(b) 给出了旋转 75° 时的图像)。识别结果如表 1 所示。从表 1 中可以看出, 在 90° 倍数的旋转下分类识别率为 100%, 最低识别率也有 82.5% (旋转角度为 $15^\circ, 150^\circ, 345^\circ$ 时)。可见, 这种分类方法对旋转变化具有一定的容忍性。



(a) JPEG 图像



(b) 旋转了 75° 的纹理图像

图 6 10 类纹理图像

本文提出的压缩域纹理图像分类方法具有一定的抗旋转能力, 并且直接在 RLE 码流中进行处理计算, 对 DCT 压缩图像只需进行熵解码, 要处理的数据量也较少, 大大缩短了处理时间。在 DCT 域中, 本文合理组织了频率系数, 使其具有与三级小波分解相似的多分辨率结构, 这种结构有利于压缩域的纹理分类及其它图像处理。在纹理分类的算法设计中考虑到同一分辨率下的不同方向区域的综合特征, 降低了分类算法对方向性的敏感度, 因此具有一定的抗旋转能力。实验结果表明这种基于 DCT 压缩域的纹理图像分类方法具有较高的正确性。下一步的工作将主要利用 DCT 域所具有的这种多分辨率特征进行基于内容的压缩域图像检索技术研究。

表 1 各旋转角度下 10 类纹理图像的分类识别率

旋转角 (\circ)	0	90	180	270	15	30	45	60	75	105	120	135
识别率 (%)	100				82.5	87.5	90	85	85	85	85	90
旋转角 (\circ)	150	165	195	210	225	240	255	285	300	315	330	345
识别率 (%)	82.5	85	85	87.5	90	85	82.5	85	85	90	85	82.5

参 考 文 献

- [1] 张皓, 侯杰昌, 廖孟扬, 基于“势”最大化的纹理分类方法, 电子科学学刊, 1999, 21(3), 315-319.
- [2] 徐朝伦, 王晓湘等, 基于子波变换的纹理图像分类, 电子科学学刊, 1999, 21(3), 404-407.
- [3] T. Chang, C. J. Kao, Texture analysis and classification with tree-structured wavelet transform, IEEE Trans. on Image Processing, 1993, 2(4), 429-441.
- [4] Shih-Fu Chang, Compressed-domain techniques for image/video indexing and manipulation, IEEE ICIP'95 Special Issue on Intelligent Signal Processing, Washington DC, USA, 1995, 314-317.
- [5] J. R. Smith, S. F. Chang, Transform features for texture classification and discrimination in large image databases, In Proc. ICIP, Austin, USA, 1994, 3, 407-411.
- [6] B. C. Smith, L. A. Rowe, Algorithms for manipulating compressed images, IEEE Computer Graphics and Applications, 1993, 13(5), 34-42.
- [7] Shih-Fu Chang, Some new algorithms for processing images in the transform compressed domain, Proc. SPIE Visual Comm. and Image Processing, Bellingham, WA USA, 1995, 445-454.
- [8] 黄祥林, 沈兰荪, 基于 DCT 压缩域的纹理图像分类, 电子与信息学报, 2002, 24(2), 216-221.

TEXTURE-IMAGE CLASSIFICATION WITH
ROTATION-INVARIANCE IN COMPRESSED-DOMAIN

Huang Xianglin Shen Lansun

(Signal and Info. Processing Lab, Beijing Polytechnic University, Beijing 100022, China)

Abstract A novel method of texture-image classification with rotation-invariance in DCT compressed-domain is proposed. By reordered the DCT coefficients for multi-resolution, the textural features are extracted in DCT transformed-domain and these textural features can be used to classify texture-images in RLE bit-stream directly with rotation-invariance. The experimental results demonstrate that texture-images can be classified at higher correct classification rate.

Key words Image processing in compressed-domain, DCT, Texture-image classification, Rotation-invariance

黄祥林: 男, 1967年生, 博士, 讲师, 研究方向: 智能化信息处理.

沈兰荪: 男, 1938年生, 教授, 博士生导师, 研究领域: 图像编码与传输, VLSI 实时信号处理, 谱信号检测等.