

基于压缩域的图像检索技术

李晓华 沈兰荪

(北京工业大学信号与信息处理研究室 北京 100022)

摘 要 图像检索技术是多媒体应用中的关键技术. 现有的基于内容图像检索技术大都是基于非压缩域的. 对于目前普遍存在的压缩格式图像, 采用这种技术必须先解压再检索, 不但计算量大, 而且需占用较多的中介存储空间, 所以严重影响了检索系统的实时性和灵活性. 同时各种压缩标准(如 JPEG, MPEG, JPEG2000 等)的推出与普及也促使人们寻求可以直接在压缩域操作的检索技术. 该文对现有的压缩域图像检索技术的发展进行综述, 并讨论了未来可能的研究方向.

关键词 基于内容图像检索; 压缩域检索; 压缩标准; 小波变换; 矢量量化

中图法分类号 TP391

Image Retrieval Technique in the Compressed Domain: A Review

LI Xiao-Hua SHEN Lan-Sun

(Signal & Information Processing Laboratory, Beijing Polytechnic University, Beijing 100022)

Abstract Image retrieval techniques are crucial in multimedia applications. A number of retrieval techniques that operate in the pixel domain have been reported in the literature. However, the large volumes of visual data necessitate the use of compression techniques. Hence, the visual data in future multimedia databases is expected to be stored in the compressed form. In order to obviate the need to decompress the image data and then apply traditional pixel domain retrieval techniques, it is efficient to index the image in the compressed domain. At the same time, the advent of compression standards has also led to the proliferation of retrieval techniques in the compressed domain. Comparing with traditional decompress-process approach which applies the pixel domain retrieval techniques, these techniques have a lower cost for computing and storing the indices. This paper presents a critical review on the compressed domain retrieval techniques proposed in the literatures. In addition, the future promising directions are also discussed.

Keywords content-based image retrieval; compressed domain retrieval; compression standard; wavelet transform; vector quantization

1 引 言

随着信息技术的发展, 图像检索受到广泛的关注. 对于非结构化的图像数据, 传统的基于文字的检索方法效率很低, 为此人们提出了基于内容的检索技术, 并且取得了一定的进展^[1~4]. 但是这些检索系

统中采用的技术(见图 1(a))大多是基于像素域的, 也就是说检索所基于的特征均从原始像素域抽取.

近年来出现了各种压缩标准如 JPEG 系列、MPEG 系列以及 H26X 系列等. 压缩标准的运用减少了存储和传输的开销, 却增加了相应的处理开销, 因为按照像素域处理方法, 这些数据必需解压后才能进行处理, 而解压的计算量是很大的. 解决这一问

题的途径之一是在压缩域直接处理图像数据,以完成传统的只有在像素域才可进行的处理,即压缩域检索技术(见图 1(b)).这样做有两方面的优势:(1)省了解压过程,处理速度快;(2)由于直接操作压

缩数据,处理过程中资源占用少.

下面,本文对现有的压缩域图像检索技术进行评述.

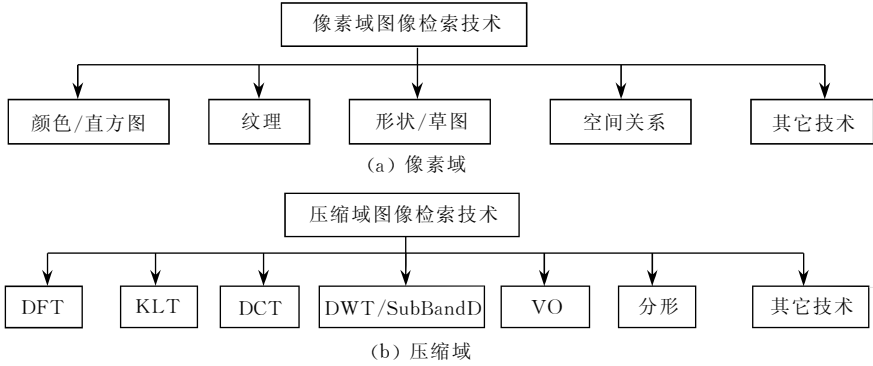


图 1 基于内容的图像检索技术

2 压缩域图像检索

压缩域的图像检索是指通过挖掘图像压缩时的中间结果或最终码流中包含的信息,力争在不解码或部分解码的情况下提取表征图像内容的特征,并以其作为索引而实现基于内容的图像检索.基于压缩域的图像检索最理想的情况当然是直接在最终的

压缩码流上(如图 2 位置 1)提取特征,但是由于熵编码属于非结构化、非字节对齐编码,在熵编码后进行这些操作十分困难,通常都在熵解码后(如图 2 位置 2)或熵解码和反量化后(如图 2 位置 3)进行.压缩域检索根据所基于的压缩方法的不同可分为两大类:变换域技术和空域技术.变换域技术包括基于 DFT, KLT, DCT, 子带/DWT 等,空域技术有基于矢量量化和分形等.

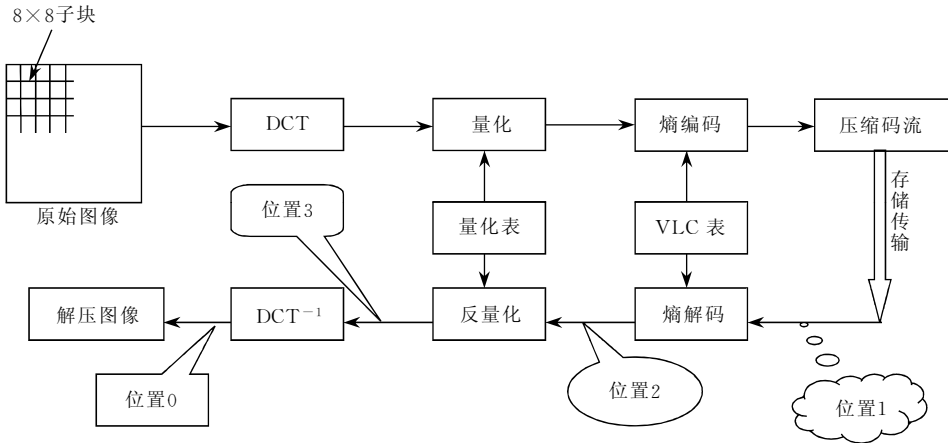


图 2 JPEG 基本模式编解码过程

2.1 离散傅立叶变换域(DFT)

DFT 变换是图像处理中的一个重要的数学变换,在图像压缩编码中主要利用它所具有的能量集中性. DFT 被用于图像检索,主要是因为它的以下两个特性: DFT 系数幅值的平移不变性; DFT 系数可用来有效地计算图像的空域相关性.

Stone 等人提出一种在 DFT 域的图像检索算法^[5]. 该算法采用两个域值来独立地调整匹配的准

确性. 用一个阈值来控制亮度的匹配,用另一个阈值来控制图像纹理的相似程度,用这两个阈值可以达到由粗到精的渐进匹配. Celantano 等研究了基于 DFT 系数角度分布的图像检索方法^[6]. 在 180 度范围扫描 DFT 谱图,统计每一角度范围的中频 DFT 系数之和,得到一角度分布直方图,以此直方图作为特征矢量进行图像检索. 虽然该特征矢量具有平移不变性,但是在图像发生旋转时,该特征矢量将发生

循环移动,因而该方法不具有旋转不变性,对于这一缺点,可通过对特征矢量进一步做 FFT 变换使其具有旋转不变性. Augustejin 等提出利用 DFT 系数的统计特征和 DFT 系数径向分布结合角度分布实现卫星图像的纹理分类^[7]. 其中的统计特征有最大幅度、平均幅度、幅度能量和幅度方差等;DFT 系数的径向分布在一定程度上反映了图像纹理的粗糙度,而 DFT 系数的角度分布则反映了图像纹理的方向性. 另外,作者在该文中还对比了两种方法的纹理分类性能,结果表明当图像中存在明显突出的频率成分时,基于 DFT 系数径向及角度分布的方法具有较好的分类效果,而当频率成分比较均匀时,基于统计特征的方法则具有更好的分类效果.

虽然基于 DFT 压缩域的检索技术取得了较好的效果,但是由于 DFT 是一种复数变换,运算量大,实用困难. 另外对于一般的图像信号,利用 DFT 进行压缩效果也不很理想,因此很少被用于图像压缩,从而使得基于 DFT 压缩域的检索能力也就无法很好体现. 目前虽然仍有学者在做这方面的研究^[8],但只是把 DFT 作为提取特征的一种技术,不再强调压缩域检索这一概念.

2.2 K-L 变换域

K-L 变换具有最好的能量集中性,是一种统计最优变换. KLT 的基本变换阵是随图像自适应变化的,所以可以通过映射图像到 K-L 空间,然后比较 KLT 系数来实现图像检索.

Pentland 等提出了一种基于 KLT 系数的人脸识别技术^[9]. 首先,随机抽取一组人脸图像,根据这组图像计算最优变换矩阵;然后利用此变换矩阵计算查询图像和目标图像的 KLT 系数;最后计算查询图像和目标图像的 KLT 系数之间的欧式距离,并依据此距离进行人脸识别. 因为 KLT 系数中只有最前面的系数较大,所以图像的特征可以用前面的少数系数(15~20)来表征,这样就可以降低特征矢量的维数,减少图像匹配计算量. 虽然 K-L 变换能够抽取一幅图像的最主要特征,即本征特征,然而这些本征特征中常常携带一些与图像检索或目标识别无关的信息,如光照方向、光照强度等,即使增大本征特征的数量也无法消除这些无关信息. 针对这一问题,Swets 等提出了一种区别 K-L 变换^[10] (Discriminant K-L Transform, DKLT),主要思路是对图像先进行 K-L 变换,然后对 K-L 变换的结果再进行区别分析,从而得到一组最具有区别能力的特征 (Most Discriminating Features, MDFs). 在 DKLT 中,不同类之间的距离被最大化,而同类之间的距离

被最小化,作者的实验结果表明 DKL 方法在检索效率方面要好于 KLT 10%~30%. 但是这一检索效率的提高是以高额的计算量为代价的,我们知道 KLT 的计算复杂度本身就很高,而 DKLT 的计算量几乎是 KLT 的 2 倍,另外还需要人工标注每幅训练图像的类别,显然,这种方法仅适合一些实时性要求不高的场合,如人脸识别等,而对于网上查询、大容量图像数据库检索等是不适合的.

到目前为止,由于其固有的计算复杂性,KLT 很少被用于传统图像压缩. 然而,K-L 变换常常被用于多波段图像分析处理和编码中^[11],因此,在遥感图像分析与检索中具有一定的潜能.

2.3 离散余弦变换域(DCT)

离散余弦变换(DCT)作为次最佳的正交变换,因其所具有的良好解相关特性和快速算法而被国际压缩标准如 JPEG, MPEG1/MPEG2/MPEG4, H261/H263 等普遍采用. 图 2 给出了基于 DCT 的 JPEG 基本模式编解码算法示意图,该算法具有普遍性,上面提到的其它国际标准几乎均采用与它相似的算法结构. 在 JPEG 中,压缩分 3 个步骤: DCT 计算、量化、熵编码. 首先,原始图像被分割成 8×8 不交迭的子块,然后对每一子块进行 2D DCT 计算并根据 JPEG 标准提供的量化表进行量化. 这样每一个子块将得到 1 个 DC 系数和 63 个 AC 系数,对所有这些系数进行 zigzag 排序,得到一个 1D 序列,所有 DC 系数采用 DPCM 进行编码,而 AC 系数采用游程长加 Huffman 技术进行编码. 解码算法正好是编码的反过程. 从整个的编码过程,我们可以看到这种基于 DCT 的标准并未考虑图像检索功能. 基于 DCT 压缩域的检索技术就是在现有的压缩标准基础上,通过分析 DCT 变换系数以及压缩算法的特点,力争在不解码(图 2 位置 1)或部分解码(图 2 位置 2、位置 3)的情况下实现图像检索. 图 2 中位置 0 是像素域检索技术的操作位置. 下面介绍并讨论各种基于 DCT 压缩域的检索技术,这些方法大多是在位置 2 或位置 3 处进行操作的.

基于 DCT 压缩域的检索技术中研究最多的是纹理特征的提取及基于纹理特征的检索和分类. Smith 等首先提出一种基于 DCT 的检索方法^[12]. 该算法将图像分成 4×4 的子块进行 DCT 变换,每一个子块得到 16 个变换系数,对图像中所有 DCT 变换块中同一位置的系数计算其均值和方差,形成一个具有 32 个分量的特征矢量,用它表征整个图像的纹理特征. 同时作者还采用了 Fisher 区别分析 (FDA) 技术降低特征矢量的维数,降维后的特征矢

量被用于支持检索. 严格地讲, 这种方法不属于压缩域检索方法, 因为它采用的 DCT 块与标准不兼容, 要采用此方法, 必须完全解码标准压缩图像, 但是此方法给我们提供了一种基于 DCT 压缩域进行图像检索的思路. Reeves 等人在此基础上进行改进, 实现了基于 DCT 图像纹理特征的提取^[13]. 他采用与标准兼容的 8×8 的子块, 并且他认为反映不同纹理特性的主要是前面的几个 AC 系数, 所以只对每个变换块内的前 8 个 AC 系数计算其方差用来形成特征矢量, 这样即使不使用文献[12]中采用的 FDA 降维方法, 维数也很小, 匹配复杂度大大降低. 文献[14,15]中介绍了另外两种基于 DCT 的图像检索方法. 与前面两种计算单个通道均值和方差的策略不同, 文献[14]根据 DCT 系数的特性, 把 DCT 系数分成不同的区域, 通过统计这些区域的系数能量, 得到一个具有 9 个分量的综合特征, 其中 4 个反映频率特性, 5 个反映空间特性. Sim 等在文献[15]中引入了人眼视觉特性, 利用一个符合人眼视觉的掩模对各个频率通道的特征分量做加权处理. Lay 等考虑到直方图技术在图像检索中的作用, 提出了一种基于低频 DCT 系数能量直方图的检索方法^[16]. 该方法首先选择 DCT 块中一个对称区域, 然后统计该区域中 DCT 系数的能量直方图, 并以此直方图为索引进行图像检索. 我们注意到, 当该区域仅选择 DC 系数时, 得到的直方图即 DC 图的灰度直方图, 而随着区域的扩大, 直方图中将包含进各种频率成分. 作者在文中建议选择 DC 系数和前 3 个 AC 系数构建直方图. Huang 等通过重组 DCT 系数, 实现了一种类似于小波分解的纹理特征提取方法^[17].

我们通过对上面的基于通道能量的纹理特征提取方法进行模拟实现, 发现这些纹理特征的组织方法抗干扰能力通常不是很强, 当图像由于压缩倍数较高或有噪情况下, 检索及分类效果往往不好. 为此我们实验室黄祥林等提出一种基于 DCT 区域能量方向性和多分辨率特征组织的纹理分类方法^[18]. 因为一定区域的变换系数代表着一定方向的频谱成分, 所以把 8×8 子块 DCT 变换系数分成直流区、竖直纹理区、水平纹理区和对角纹理区, 并且直接从 RLE 码流中计算这些区域的能量, 通过一种多分辨率特征组织形成特征矢量来支持纹理分类及检索, 实验结果表明这种方法不仅计算简单, 而且在有噪情况下也具有很好的分类结果. 黄祥林等还提出了一种基于 DCT 频率系数分布特点的纹理分类方法^[19], 对 8×8 子块 DCT 变换系数按低频率、中频率、高频率分区, 从 RLE 码流中直接统计这些区域

的能量, 形成多分辨率特征矢量. 由于这里是把各个方向中同等地位的系数考虑在同一个区域, 因此这种方法具有较好的旋转不变性. 这两种方法的另外一个特点是结合 JPEG 压缩算法的具体特点, 直接从 RLE 码流中计算特征矢量, 真正实现了基于压缩域的图像检索.

除了纹理特征, 边缘信息在表征图像内容上也具有重要意义. Abdelmalek 和 Hershey 提出了一种利用 DCT 系数检测方向线的技术^[20]. 这种技术基于这样一种现象: 如果原始图像在水平、垂直、对角方向具有明显的边缘特征, 那么其 DCT 系数将相应地在垂直、水平、对角方向拥有较大的值. 作者指出在空域斜率为 m 的直线在 DCT 域中表现为斜率大约为 $1/m$ 的直线, 这就为直接应用空域中的边缘检测技术到 DCT 域提供了理论依据, 对于大量的基于 DCT 的压缩图像, 进行边缘检测时就可以省去 DCT 反变换这一步, 从而大大降低处理时间. Shen 等人研究了一种利用高频 DCT 系数检测感兴趣区及其边缘的算法^[21]. 实验结果表明该算法的边缘检测结果达到了对原始图像采用 Sobel 边缘检测算子的水平. Lee 等人提出了一种利用边缘信息进行图像匹配的方法^[22], 这种方法包括 3 个步骤: 首先, 从 AC 系数中提取二值边缘图; 然后, 利用边缘图计算边缘方向、边缘强度和边缘偏移量形成特征矢量; 最后按一定的相似性准则进行图像匹配, 实验结果表明该方法具有较好的边缘检测能力和检索效率. 黄祥林等也提出一种 DCT 域基于边缘轮廓和连通直方图的检索方法^[23], 通过重组 DCT 系数获得图像的多级轮廓, 以各级轮廓的连通直方图加权 and 作为特征量进行图像检索. 实验结果表明该方法具有很好的检索效果.

除了上述的纹理特征和边缘信息的提取, 研究者还提出了其他的基于 DCT 域的图像检索方法, 这些方法可概括地称之为基于统计特征的检索方法. Shneier 等人提出了一种基于 JPEG 压缩域的图像检索算法^[24]. 它的主要思想是通过判定检索图像和目标图像中不相连区域对中 DCT 系数的相似关系的大小来进行检索: 首先在图像内选取 $2K$ 个互不相连的区域窗, 随机配对得到 K 个窗对, 对每个窗计算其包含的 8×8 子块中每个系数的均值得到一个 64 维的特征矢量, 将配对窗的特征矢量之间按其对应关系判定每一对分量的相似度大小并赋予该窗对一个比特(0 或 1), 而检索图像和目标图像的相似与否取决于所有这些窗对比特流的相似度. Yu 提出一种可用于直接比较两幅 JPEG 压缩图像相似

度的测度:Q-测度,这种测度可以直接利用 DCT 系数计算两幅图像之间的相似程度^[25].文献[26]介绍了一种直接从 JPEG 图像的 RLE 码流中提取特征矢量进行检索的方法.主要思路是应用主分量法到 DCT 系数,把 DCT 图像映射到一个低维空间,用一个低维特征量(特征空间一点)来表征图像.另外该文还分析了量化对检索结果的影响,认为量化不会造成特征映射的大畸变.这些方法虽然取得了一定的效果,但是总体来说,它们一般没有具体理论依据,并且计算量较大,可以说不属于基于 DCT 压缩域的主流检索技术.

2.4 离散小波变换域(DWT)

由于小波变换具有良好的时频局部性及与人眼视觉特性相符的多分辨率分析能力,因而它一经出现就被广泛地应用于图像压缩领域.近几年的研究表明基于小波变换的静止图像压缩系统在性能上优于基于 DCT 的压缩系统,从而促使小波变换在 JPEG2000 和 MPEG-4 中(静态纹理的编码)获得应用.目前小波域图像检索技术已得到多方面的研究.

Smith 等人提出了基于小波系数的纹理识别算法^[27].在他们的方法中,对图像进行三级小波分解并计算所有高频子带小波系数的幅度值,所有幅度系数子带通过上采样恢复到与原始图像同样的尺寸,这样一来共得到 9 个纹理通道,所有通道中位置相同的点构成一个 9 维矢量,对于 $M \times N$ 大小的图像,将得到 $M \times N$ 个 9 维矢量,对每一矢量中的分量进行二值量化处理,处理后的矢量共有 $2^9 = 512$ 种可能值,从而可构建 512 级统计直方图,并用该直方图作为纹理特征来支持图像的检索.虽然这种方法具有较好的检索性,但是计算复杂度较高.Mandal 等在此基础上提出了一种改进算法——快速小波直方图检索方法^[28],在基本不影响检索性能的情况下,大大减低了计算复杂度.这一类方法被统称为小波直方图法,是 DWT 压缩域的典型算法之一,不但对纹理图像具有较好的检索效果,而且对于自然图像也很有效.

基于 DWT 压缩域的另一类典型方法是子带能量法.这种方法的基本思路是对原始图像做小波分解,得到多个时频子带,计算每个子带的能量形成多维特征矢量用于检索.Chang 等人提出的基于非规则树分解的纹理分析方法^[29]是子带能量法的最早雏形,在该方法中,作者通过计算 J 个中分辨率子带系数的能量形成 J 维特征矢量用于纹理匹配,取得了较好的纹理分类结果.Lee 等在文献[30]中为了降低小波包分解时能量特征矢量的维数,仅仅选

择包分解子带中能量较大的 7 个子带形成特征矢量用于纹理分类,试验表明在降低计算复杂度的同时,也可以取得很好的分类结果.近年来,很多研究者通过选取不同的小波基(正交基、双正交基)、不同的小波分解方法(塔式小波分解、小波包分解等)以及不同的子带能量计算法等对这种方法进行了深入研究.文献[31]比较了不同小波变换用于纹理分类的优劣性,主要比较了传统塔式正交小波变换、塔式双正交小波变换、树结构正交小波分解、树结构双正交小波分解、Gabor 小波变换五种情况.文中采用各子带能量分布的前三阶矩构建特征矢量,实验结果表明,正交小波变换取得比双正交小波变换稍好的分类效果;树结构小波分解虽然特征矢量具有更高的维数且计算量大,但并不比传统塔式小波变换效果好很多;在所有的结果中 Gabor 小波变换具有最好的纹理分类结果,但计算量也最大.Chang 等在文献[32]中比较了子带/小波域和 DCT 域的纹理分类性能.该文中以各子带能量作为纹理特征.实验结果如下:子带(92%),小波(92%), 4×4 DCT(85%),可以明显看到,同 DCT 域相比,DWT 域具有更好的纹理分类性能,这是因为 DWT 的能量集中性要好于 DCT.

跟上面的基于小波子带能量的方法不同,Mandal 等人把小波分解的方向子带统计直方图特征用于图像检索^[33].他们认为对原始图像而言,不同的图像可能具有相同的统计直方图特性,但不同图像小波分解方向子带的统计特征不可能完全相同.然而直接进行所有小波分解子带直方图的比较计算复杂度过高,采用直方图的矩特征能大大降低特征数据量.因此他们构造了不同子带的矩特征用于支持小波域的图像检索.同时,由于小波系数的分布近似符合广义高斯分布,Mandal 又提出对不同子带统计特性的匹配只需比较其用于描述分布的标准方差和形状参数,从而实现了一种低复杂度高效率的图像检索算法.

Sun 等证明了任何一种对称小波基都可以作为边缘检测算子来实现图像的边缘检测^[34],只是采用反对称小波基时,检测出的边缘图像会产生半个像素的平移.这为 DWT 压缩域基于边缘轮廓的检索鉴定了理论基础.我们可以直接从图像的小波变换获取图像的多分辨率边缘图像,并进而基于这些边缘图像实现图像的检索.魏海等提出了一种基于反对称双正交小波变换的图像检索方法^[35].首先,对图像进行反对称双正交小波变换,然后,在变换域提取多尺度边缘,最后利用边缘像素的方向梯度相角

信息构建直方图统计特征量以支持检索. 实验结果表明该方法不仅具有较高的检索效率, 同时具有较强的光照变化鲁棒性和一定程度的抗几何变化(尺度、平移、旋转等)的能力. 文献[36]也实现了一种基于小波和不变矩的形状检索方法. 对原始图像进行小波变换, 求取多尺度边界图像; 计算每一边界图像的 7 个不变矩形成特征矢量并归一化; 计算查询图像与目标图像特征矢量之间的距离来确定他们之间的相似性. 文中试验表明该方法能较好地描述图像的形状及空间分布信息, 并且具有较好的平移、尺度、旋转不变性. 另外, Loupias 等利用小波变换所具有的时频局域性, 进行特征点检测, 然后基于这些感兴趣的特征点提取纹理、颜色特征实现图像检索^[37,38].

Jacobs 等人基于小波变换系数的直接比较实现了一种快速的多分辨率图像检索算法^[39]. 在该算法中所有图像的尺寸都调整为 128×128 并进行小波分解, 对每一图像选出幅值最大的 $M(40 \sim 60)$ 个小波系数记录其颜色、符号以及索引位置, 将这些数据组织起来作为表征图像的特征用来支持检索. 虽然该文献给出了较好的检索性能, 但由于特征数据中包含系数的位置索引, 所以该算法明显不具备对平移和旋转等几何变化的鲁棒性. Wang 等人提出了另一种基于小波分解系数比较的图像检索算法^[40]. 所有图像也调整为 128×128 的尺寸并进行四级小波分解, 利用最低一级的 4 个 8×8 子带图像(LL, LH, HL, HH)进行三级检索, 第一级通过 LL 子带系数方差的比较筛选出 20% 的图像; 在第二级中对上一级筛选出的图像进行 LL 子带小波系数的直接比较; 最后一级通过 LL, LH, HL, HH 四个子带图像的系数的比较得到最终的检索输出图像. 虽然该文献给出了比文献[39]更好的检索性能, 然而该算法同样也不具备对平移和旋转的鲁棒性.

文献[41]提出了一种利用各子带之间的相关性进行图像检索的方法. 首先做小波分解, 分别用各级的 LL 子带, $LL + HL$ 子带, $LL + LH$ 子带, $LL + HH$ 子带近似重构上一分辨率图像, 并计算它们与用 $LL + HL + LH + HH$ 精确重构的图像之间的相关性, 在每一层形成一具有 4 个分量的特征矢量, 用来进行图像分级检索.

除了上面这些直接基于小波变换系数提取各种表征图像内容特征的方法外, 最近不少研究者们开始研究是否能直接从压缩码流中提取表征图像内容的特征. 因为小波零树编码方法被公认为是当前最好的基于小波变换的图像编码方法, JPEG2000 中

的熵编码部分就采用了零树编码的思想, 所以针对小波零树编码算法研究者提出一些简单有效的检索技术. Wilson 等通过统计 EZW 每一编码层的重要系数数目, 并结合各层的阈值来计算各子带能量^[42], 这种方法不需要重建和存储整个小波分解矩阵, 而是一边扫描压缩符号流一边计算特征矢量, 理论分析和实验结果表明这种方法可以在大大减少计算量和资源需求量的基础上得到较准确的子带能量特征. 随着 JPEG2000 压缩标准的推出和普及, 这种基于压缩符号流的检索技术将会得到更广泛的研究.

2.5 矢量化(VQ)

矢量化是一种高效的图像压缩方法, VQ 将图像中的 N 个样点合为一组, 形成 N 维空间中的一个矢量 \mathbf{R}^K , 映射每一矢量 \mathbf{R}^K 到一个有限的码本集合 S 中的某一码字矢量 \mathbf{X}_i , 用 \mathbf{X}_i 的索引值 i 来表征 \mathbf{R}^K , 这样就达到压缩的目的. 由此可见, 矢量化本身就是一种索引机制, 可以用来进行图像检索.

Idris 等利用矢量量化的压缩数据——码字索引集合, 构造了码书中码字矢量使用频数的统计直方图, 将其作为表征图像内容的特征量用于比较检索图像和待检索数据库中图像的相似程度^[43]. 同时, 为了进一步降低直方图特征量的存储要求及其相互间距离计算的复杂度, 他们还还对码字使用频数直方图进行了简化处理, 提出了使用图(usage map)特征量, 大大减小了特征量的存储量和距离计算的复杂度, 其代价是检索正确率的降低. 若根据图像中矢量的特性, 首先进行矢量分类, 比如根据矢量方差值, 把矢量分成平滑矢量(方差较小)和细节矢量(方差较大), 不同类矢量对应不同的码本. 矢量化时根据矢量的类型到相应的码本中寻找其码字索引, 这种矢量化称为分类矢量化. 基于分类矢量量化的检索方法与上述的基于矢量化类似, 不同之处仅在于分别统计各类码本中码字矢量使用频数的统计直方图, 然后把他们连成一个统一的直方图. 基于分类矢量化进行检索有一个优点: 可以根据人眼对各类矢量的敏感程度, 对直方图中的分量采用不同的权值来提高检索效率. 魏海等比较了标量量化(SQ)、矢量量化(VQ)和分类矢量化(CVQ)用于图像检索的性能^[44]. 结果表明, VQ 的检索性能优于 SQ, 而 CVQ 的检索性能明显优于 VQ.

2.6 分形/仿射变换

迭代分形编码作为一种高效的图像压缩方法引起了广泛的关注. 图像经迭代分形编码后的结果——分形码中所包含的迭代函数系统以一种紧凑的形式唯一表征了原始图像, 因此可以利用迭代函

数系统的某些特性来实现基于迭代分形压缩数据的图像检索操作. Zhang 等人最早提出一种直接利用图像的分形码来实现基于纹理的图像检索技术^[45], 这里分形码是以图像子块为单位得到的, 即首先把图像分成不重叠的若干子块, 对每一子块根据分形原理寻找其分形码. 检索时直接比较查询图像和目标图像的分形码. 在文献^[46]中, Zhang 等进一步比较了基于分形的方法与小波子带能量法的检索效果. 结果表明对于纹理图像, 小波方法要好些, 而对于其他图像, 则分形方法性能略胜一筹.

2.7 混合压缩域

图像压缩的混合技术是指结合两种以上基本编码方法的压缩技术. 这种结合充分利用相应基本方法的优势, 从而可以取得更优越的性能, 比如在高效压缩的同时, 可以直接支持检索操作. Muneesawang 等提出了一种小波域基于 VQ 的检索技术^[47]. 首先对图像作小波分解, 然后对多分辨率小波系数进行矢量量化, 最后构建码书中码字矢量使用频数的统计直方图作为特征量支持检索. 魏海等提出了一种小波域基于 CVQ 的检索技术^[48]. 由于分类矢量量化将图像中不同类别的矢量分开, 用各自独立的码书分别进行量化, 和采用单一有限尺寸码书的矢量量化方法相比, 不但能够实现高效压缩, 而且具有更好的检索性能. 魏海等还提出了一种小波域基于分形的检索技术^[49]. 在小波域对高频系数进行迭代分形编码, 形成分形码, 然后直接用这些分形码作为特征量实现图像检索.

3 讨 论

上面介绍了在各种压缩域的图像检索技术, 我们很难说这些技术孰优孰劣. 一种检索技术是否能得到普及应用很大程度上取决于其所基于的压缩方法的性能. KLT 虽然统计最优, 但其固有的计算复杂性限制了它及基于它的检索技术的广泛应用. DFT 在图像分析中具有重要的作用, 但是 DFT 对非周期的图像信号, 其压缩能力往往不够理想, 很少被用于图像压缩, 因此其检索能力也就无法很好体现. 矢量量化的编码过程本身就是一种索引机制, 最容易实现压缩域图像检索, 但矢量量化的编/解码过程很不对称. 分形虽然在编码方面具有很大的潜能, 但是分形码是高度非线性并且与待编码图像关系很大, 只有图像本身具有明显的自相似性或统计相似性时, 才能获得较高的压缩效率, 另外分形的编/解码过程也很不对称. 由于这些原因, 基于 DFT, KLT, 矢

量量化和分形等的图像检索技术在做了探索性的研究后, 就不再成为基于压缩域检索技术的研究主流. 但另一方面, 在一些特殊领域, 如多波段遥感图像、医学图像中, 由于常常要根据这些图像的特点选用最合适的压缩方法, 这些检索技术仍然有他们的潜力所在.

相对而言, 由于目前的编码标准中普遍采用的是基于 DCT 变换和小波变换的压缩算法, 就使得基于这两种压缩域的检索技术研究异常活跃. 在这些研究中, 基于 DCT 压缩域和基于小波压缩域均表现出较好的检索性能. 然而相对来说, 由于小波变换所具有的以下优势, 使得基于小波变换的检索技术将更有前途: (1) 多分辨率分析能力; (2) 明显的方向信息; (3) 简单而有效的边界及形状检测能力; (4) 与 VQ、分形等技术的易结合能力. 并且随着以小波变换为核心技术的 JPEG2000 标准的推出与普及, 基于小波变换的检索技术将会有更广泛的应用前景.

压缩域图像检索的另一个发展方向是从图像压缩的角度进行压缩域图像检索技术的研究, 即把压缩与检索结合起来, 研究支持检索的图像压缩方法. 如文献^[47~49]就是从这样的思路出发, 研究能同时实现高效压缩和有效检索的技术.

本文主要介绍基于压缩域图像检索技术中的图像处理部分, 即特征提取技术, 然而, 特征之间的相似性匹配对于检索效率的提高也有很重要的意义. 如何针对具体的特征矢量, 利用统计学理论, 寻找与各种压缩域特征矢量最相符的相似性匹配方法这两年得到越来越多的关注^[50,51].

上面介绍的技术的操作位置大部分是处于图 2 的位置 3 处, 也就是说在提取特征量以前, 首先需要重建 DCT 和 DWT 变换系数并存储他们, 然后才能在此基础上提取特征量. 为了更进一步减少计算量和资源消耗量, 依据文献^[42]的思路, 结合具体压缩算法的特点, 研究直接从压缩码流或熵解码后的符号流中提取特征的方法将具有很重要的实用意义, 也将是压缩域检索技术的一个发展方向. 目前的研究方法均是在一定的约束条件下, 如图像大小一定、压缩比一定、小波分解级数一定等, 但是实际情况往往不是这样, 人们面对的往往是各种大小、各种质量、各种压缩参数下的图像, 因此如何实现对形形色色图像的鲁棒检索将是基于压缩域检索技术的一个难点. 压缩域检索技术的另一个难点也是基于内容检索技术的难点, 即如何实现低级物理特征和高级语义特征之间的映射, 从而实现直接语义特征的压

缩域图像检索.

4 结 语

本文简要介绍了目前国内外压缩域图像检索技术的发展动态及其研究内容与方法. 虽然像素域图像检索技术已经有一些公认的、成熟的技术(如 MPEG7 中定义的各种内容描述子), 但是压缩域检索技术因压缩算法的多样性、压缩域操作的复杂性, 仍然处于初步研究阶段, 各种尝试性的算法层出不穷. 我们相信, 随着 Internet 的发展及压缩标准的不断推出, 压缩域图像检索技术将会在基于内容图像检索中受到越来越多的重视, 并将成为图像检索领域的一个新的研究热点.

参 考 文 献

- Flickner M *et al.* Query by image and video content; The QBIC system. *IEEE Computer Magazine*, 1995, 28(9): 23~32
- Bach J R *et al.* The virage image search engine: An open framework for image management. In: *Proceedings of SPIE: Storage and Retrieval for Still Image and Video Databases*, San Jose, 1996. 76~87
- Pentland A *et al.* Photobook: Content-based manipulation of image databases. *International Journal of Computer Vision*, 1996, 18(3): 233~254
- Smith J R, Chang S-F. VisualSEEK: A fully automated content-based image query system. In: *Proceedings of the ACM Multimedia Conference*, Boston, MA, 1996. 87~98
- Stone H S, Li C S. Image matching by means of intensity and texture matching in the Fourier domain. In: *Proceedings of SPIE: Storage and Retrieval for Still Image and Video Database*, San Jose, 1996. 337~349
- Celentano A, Lecce V D. A FFT based technique for image signature generation. In: *Proceedings of SPIE: Storage and Retrieval for Image and Video Database*, San Jose, 1997. 457~466
- Augustejn M, Clemens L E, Shaw K A. Performance evaluation of texture measure for ground cover identification in satellite images by means of a neural network classifier. *IEEE Transactions on Geoscience & Remote Sensing*, 1995, 33(3): 616~626
- Sim Dong-Gyu, Kim Hae-Kwang *et al.* Translation, scale, and rotation invariant texture descriptor for texture-based image retrieval. In: *Proceedings of IEEE International Conference on Image Processing*, Vancouver, BC, Canada, 2000. 742~745
- Pentland A *et al.* Photobook: Tools for content-based manipulation of image database. In: *Proceedings of SPIE: Storage and Retrieval for Image and Video Database*, San Jose, 1994. 34~47
- Swets D L, Weng J. Using discriminant eigenfeature for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1996, 18(8): 831~836
- Saghri J A, Tescher A G, Reagan J T. Practical transform coding of multispectral imagery. *IEEE Signal Processing Magazine*, 1995, 12(1): 398~407
- Smith J R, Chang S F. Transform features for texture classification and discrimination in large image databases. In: *Proceedings of IEEE International Conference on Image Processing*, Austin, 1994. 407~411
- Reeves R, Kubik K, Osberger W. Texture characterization of compressed aerial images using DCT coefficients. In: *Proceedings of SPIE: Storage and Retrieval for Image and Video Database*, San Jose, 1997. 398~407
- Bae H J, Jung S H. Image retrieval using texture based on DCT. In: *Proceedings of ICICS'97*, Singapore, 1997. 1065~1068
- Sim D G, Kim H K, Park R H. Fast texture description and retrieval of DCT-based compressed images. *Electronics Letters*, 2001, 37(1): 18~19
- Lay J A, Ling G. Image retrieval based on energy histograms of the low frequency DCT coefficients. In: *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Phoenix, AZ, USA, 1999. 3009~3012
- Huang Y L, Chang R F. Texture features for DCT-coded image retrieval and classification. In: *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Phoenix, AZ, USA, 1999. 3013~3016
- Huang Xiang-Lin, Shen Lan-Sun. Texture-image classification in DCT compressed-domain. *Journal of Electronics & Information Technology*, 2002, 24(2): 216~221 (in Chinese)
(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)
- Huang Xiang-Lin, Shen Lan-Sun. Texture-image classification with rotation-invariance in compressed-domain. *Journal of Electronics & Information Technology*, 2002, 24(11): 1190~1196 (in Chinese)
(黄祥林, 沈兰荪. 一种具有旋转不变性的压缩域纹理图像分类方法. *电子与信息学报*, 2002, 24(11): 1190~1196)
- Abdelmalek A A, Hershey J E. Feature cueing in the discrete cosine domain. *Journal of Electronic Imaging*, 1994, 3(1): 71~80
- Shen B, Sethi I K. Direct feature extraction from compressed images. In: *Proceedings of SPIE Storage & Retrieval for Image and Video Databases*, San Jose, 1996. 404~414
- Lee Seong-Whan, Kim Young-Min, Choi Sung Woo. Fast scene change detection using direct feature extraction from MPEG compressed videos. *IEEE Transactions on Multimedia*, 2000, 2(4): 240~254
- Huang Xiang-Lin. Research on compressed-domain image retrieval techniques [Ph D dissertation]. Beijing Polytechnic University, Beijing, 2001 (in Chinese)
(黄祥林. 基于压缩域的图像检索技术的初步研究[博士学位论文]. 北京工业大学, 北京, 2001)
- Shneier Michael, Mohamed Abdel-Mottaleb. Exploiting the JPEG compression scheme for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1996, 18(8): 849~853
- Yu Hong Heather. Visual image retrieval on compressed domain with Q-distance. In: *Proceedings of IEEE International Conference on Computational Intelligence and Multimedia Applications*, New Delhi, India, 1999. 1013~1016
- Seales W B, Yuan C J, Brown M. Efficient content extraction in compressed images. In: *Proceedings of IEEE Workshop on Content-Based Access of Image and Video Libraries*, San Juan,

- Puerto Rico, 1997. 52~58
- 27 Smith J R, Chang S F. Automated binary texture feature sets for image retrieval. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Atlanta, 1996. 2239~2242
- 28 Mandal M K, Aboulnasr T, Panchanathan S. Fast wavelet histogram techniques for image indexing. *Journal of Computer Vision and Image Understanding*, 1999, 75(1): 99~110
- 29 Chang T, Kuo C C J. Texture analysis and classification with tree-structured wavelet transform. *IEEE Transactions on Image Processing*, 1993, 2(4): 429~441
- 30 Lee Moon-Chuen, Pun Chi-Man. Texture classification using dominant wavelet packet energy features. In: Proceedings of IEEE Southwest Symposium on Image Analysis and Interpretation, Austin, TX, USA, 2000. 301~304
- 31 Ma W Y, Manjunath B S. A comparison of wavelet transform features for texture image annotation. In: Proceedings of IEEE International Conference on Image Processing, Washington, DC, USA, 1995. 256~259
- 32 Chang Shih-Fu. Compressed-domain techniques for image/video indexing and manipulation. In: Proceedings of IEEE International Conference on Image Processing, Washington, DC, USA, 1995. 314~317
- 33 Mandal M K. Wavelet based coding and indexing of images and video [Ph D dissertation]. University of Ottawa, Ottawa, Canada, 1998
- 34 Sun M, Scialabasi R J. Symmetric wavelet edge detector of the minimum length. In: Proceedings of IEEE International Conference on Image Processing, Washington, DC, USA, 1995. 177~180
- 35 Wei Hai, Shen Lan-Sun. Wavelet domain image indexing using directional gradient angular histogram. *Journal of Circuits and Systems*, 2001, 6(2): 20~24 (in Chinese)
(魏海, 沈兰荪. 小波变换域内基于方向梯度相角直方图的图像检索算法. *电路与系统学报*, 2001, 6(2): 20~24)
- 36 Yao Yu-Rong, Zhang Yu-Jin. Shape-based image retrieval using wavelet and moment. *Journal of Image and Graphics*, 2000, 5(3): 206~210 (in Chinese)
(姚玉荣, 章毓晋. 利用小波和矩进行基于形状的图像检索. *中国图像图形学报*, 2000, 5(3): 206~210)
- 37 Loupias E, Sebe N, Bres S, Jolion J M. Wavelet-based salient points for image retrieval. In: Proceedings of IEEE International Conference on Image Processing, Vancouver, BC, Canada, 2000. 518~521
- 38 Sebe N, Lew M S, Tian Q, Huang T S, Loupias E. Color indexing using wavelet-based salient points. In: Proceedings of the IEEE Workshop on Content-based Access of Image and Video Libraries, Hilton Head Island, SC, USA, 2000. 15~19
- 39 Jacobs C E *et al.* Fast multi-resolution image querying. In: ACM International conference on Computer graphics and interactive techniques, Los Angeles, 1995. 277~286
- 40 Wang J Z *et al.* Wavelet-based image indexing techniques with partial sketch retrieval capability. In: Proceedings of the 4th International Forum on Research and Technology Advances in Digital Libraries, Washington, DC, USA, 1997. 13~24
- 41 Grazia Maria *et al.* Fast retrieval on compressed images for internet applications. In: Proceedings of the 5th IEEE International Workshop on Computer Architectures for Machine Perception, Padova, Italy, 2000. 136~141
- 42 Wilson B *et al.* Compressed-domain classification of texture images. In: Proceedings of the 5th IEEE International Workshop on Computer Architectures for Machine Perception, Padova, Italy, 2000. 347~355
- 43 Idris F, Panchanathan S. Image indexing using vector quantization. In: Proceedings of SPIE: Storage and Retrieval for Image and Video Databases, San Diego, 1995. 373~380
- 44 Wei Hai, Shen Lan-Sun. Performance comparison of image indexing using quantization statistical features. *Journal of Circuits and Systems*, 2001, 6(3): 11~15 (in Chinese)
(魏海, 沈兰荪. 量化方法及其统计特征量用于图像检索的性能比较. *电路与系统学报*, 2001, 6(3): 11~15)
- 45 Zhang A, Cheng B *et al.* Approach to query-by-texture in image database. In: Proceedings of the SPIE Conference on Digital Image Storage and Archiving Systems, Philadelphia, 1995. 338~349
- 46 Zhang A, Cheng B *et al.* Comparison of wavelet transforms and fractal coding in texture-based image retrieval. In: Proceedings of the SPIE Conference on Visual Data Exploration and Analysis III, San Jose, 1996. 116~125
- 47 Muneesawang P, Guan L. Multiresolution-histogram indexing and relevance feedback learning for image retrieval. In: Proceedings of IEEE on Image Processing, Vancouver, BC, Canada, 2000. 526~529
- 48 Wei Hai, Shen Lan-Sun. Image indexing using classified vector quantization. *Acta Electronica Sinica*, 2001, 29(7): 933~936 (in Chinese)
(魏海, 沈兰荪. 基于分类矢量量化的图像压缩和检索算法. *电子学报*, 2001, 29(7): 933~936)
- 49 Wei Hai, Shen Lan-Sun. Fractal-based image storage and indexing. In: Proceedings of SPIE: Storage and Retrieval for Media Databases, San Jose, 2000. 421~429
- 50 Hill P, Canagarajah C N, Bull D R. Statistical wavelet subband modelling for texture classification. In: Proceedings of IEEE International Conference on Image Processing, Thessaloniki, Greece, 2001. 165~168
- 51 Do M N, Vetterli M. Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. *IEEE Transactions on Image Processing*, 2002, 11(2): 146~158



LI Xiao-Hua, born in 1973, Ph. D. candidate. Her research interests include image/video content analysis, content-based image retrieval and image/video compression.

SHEN Lan-Sun, born in 1938, professor, Ph. D. supervisor. His research interests include image processing, video encoding/transmission and medical image processing etc.