

融合全局和局部相位特征的彩色纹理图像检索

徐佳,曲怀敬*,张志升,王纪委,魏亚南,张汉元

(山东建筑大学 信息与电气工程学院,山东 济南 250101)

摘要:针对目前彩色纹理图像检索系统存在颜色特征提取不充分、纹理特征提取与匹配的有效性亟待提高等问题,提出一种新的基于全局和局部相位特征融合的彩色纹理图像检索方法。该方法首次将局部引力角模式(Pattern of Local Gravitational Force Angle, PLGFA)作为局部相位描述符用于彩色纹理图像检索。在该方法中,首先在适合人类视觉系统的HSV颜色空间中采用非均匀方式量化三个颜色分量,并通过构建颜色直方图充分地提取颜色特征,同时在Gabor复变换域针对V分量数据进行相对相位建模而提取全局纹理相位特征;然后,利用PLGFA描述符在RGB空间的灰度图像中提取局部纹理相位特征;最后,将这三种特征通过闭式K-L(Kullback-Leibler)距离和改进版本的Manhattan距离进行有机融合,从而实现彩色纹理图像的检索。在VisTex(full/small)、STex和ALOT四个通用的彩色纹理图像库上得到的检索实验结果表明,综合考虑平均检索率和时间消耗两个性能评价指标,该方法较其他最新方法具有较明显优势,从而验证了其有效性和可行性。

关键词:纹理图像检索;局部引力角模式;全局相位特征;局部相位特征;颜色空间

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2022)09-0056-09

doi:10.3969/j.issn.1673-629X.2022.09.009

Color Texture Image Retrieval of Fusing Global and Local Phase Feature

XU Jia, QU Huai-jing*, ZHANG Zhi-sheng, WANG Ji-wei, WEI Ya-nan, ZHANG Han-yuan

(School of Information & Electric Engineering, Shandong Jianzhu University, Jinan 250101, China)

Abstract: Aiming at the problems of insufficient color feature extraction and urgent improvement of the effectiveness of texture feature extraction and matching, a new color texture image retrieval method based on global and local phase feature fusion is proposed. For the first time, PLGFA (Pattern of Local Gravitational Force Angle) is used as a local phase descriptor for color texture image retrieval. In this method, the three color components are quantized in a non-uniform way in HSV color space suitable for the human visual system, and the color features are fully extracted by constructing a color histogram. Meanwhile, the global texture phase features are extracted by relative phase modeling of V-component data in the Gabor complex transform domain. Then, the PLGFA descriptor is used to extract local texture phase features from the grayscale image in RGB color space. Finally, the three features are organically fused with the closed-form Kullback-Leibler (K-L) distance and the improved Manhattan distance to achieve color texture image retrieval. Experimental results obtained on VisTex (full/small), STex, and ALOT color texture image databases show that the proposed method has obvious advantages over the state-of-the-art methods considering average retrieval rate and time consumption, thus verifying its effectiveness and feasibility.

Key words: texture image retrieval; pattern of local gravitational force angle; global phase feature; local phase feature; color space

0 引言

近年来,随着通信网络以及多媒体设备的迭代升级,使得网络传播速度和设备存储能力有了极大的提升,人们依靠网络可以方便地传输文字、语音和图像等内容。对于图像而言,由于它是一种可以使人快速获得大量信息的载体,所以在医疗、教育、交通、娱乐等活动中有着广泛的应用。在上述领域中,人们常常需要

从网络或者本地数据库中检索到一幅或同一类别的图像,如何快速准确地查询到目标图像成为了一项研究热点^[1]。近年来的研究表明,基于内容的图像检索(Content-Based Image Retrieval, CBIR)是一种高效的图像检索方案^[2]。它主要包括图像的特征提取与相似性测度两个部分,即对于一幅彩色图像,首先使用一定的方法,提取其颜色、纹理等特征,然后采用合适的或

收稿日期:2021-10-01

修回日期:2022-02-16

基金项目:国家自然科学基金项目(62003191);山东省自然科学基金(ZR2014FM016);山东省重大科技创新工程(2019JZZY010120)

作者简介:徐佳(1996-),男,硕士研究生,研究方向为图像检索;通讯作者:曲怀敬(1965-),男,教授,博士,研究方向为模式识别、基于多尺度多方向变换和深度学习的图像处理。

具有闭式形式的相似性测度与其他图像进行匹配,从而完成检索任务。而对于一幅方向信息丰富的彩色纹理图像而言,常常伴随着或颜色、或纹理方向、或兼具两者的变化^[3],因此如何构建一种准确、快速的检索方案目前成为了基于内容的彩色纹理图像检索的研究难点。

颜色特征是一种重要的视觉特征,在彩色图像检索中得到了广泛的应用。通常,颜色特征根据颜色直方图提取。Dey 等人^[4]提出在 RGB 颜色空间中分别量化三个通道的数据,并构成级联颜色直方图。与此相似,Verma 等人^[5]提出量化 HSV 颜色空间中的 H、S 通道的数据,并构建级联颜色直方图。Bhunia 等人^[6]通过考虑 HSV 颜色空间中 H、S 通道之间的关系,以组合的方式构建 H、S 联合直方图。此外,Agarwal 等人^[7]根据 HSV 颜色空间中两两不同通道之间的关系,量化并构建与纹理信息相融合的颜色特征。

除了广泛采用的颜色直方图特征之外,其他颜色特征也得到了应用。例如,Guo 等人^[8]在 RGB 颜色空间中以最大最小和块编码的方法提取颜色特征。Vinayak 等人^[9]在 RGB 颜色空间中使用颜色矩和颜色自相关图进行颜色特征提取。Li 等人^[10]在 RGB 颜色空间中利用 Gabor 小波在三个通道分别进行多尺度多方向的图像分析,以提取综合不同通道信息的特征。以上颜色特征的提取可以分为两类,一类是直接量化或编码颜色特征,提取诸如颜色直方图、颜色矩、颜色自相关图等特征,而另一类则是结合颜色信息与纹理信息的特征。然而,在以上采用颜色直方图提取特征的方法中,常使用均匀量化方式来量化各通道数据,并未针对性地提取颜色特征;同时,对于无数据分布的区间也进行了量化,得到了无效特征,因此特征性能较差。另一方面,在采用多尺度通道分解以及颜色矩等方法中,或方法过于复杂、或方法过于简单,并存在特征冗余或欠缺、特征性能差等不足。

纹理特征分为全局的和局部的纹理特征,一般而言,在空域或者变换域对整幅纹理图像进行统计建模并提取的特征为全局纹理特征。Tai^[11]基于二维 Gabor 小波的多尺度多方向变换,对图像进行各子带的全局纹理特征提取。Qu 等人^[12]在双数复小波变换域进行多尺度多方向的图像分析,并根据不同的分布模型对各子带进行全局纹理特征提取。此外,Etemad 等人^[13]使用 Shearlet 变换分别在 RGB 图像的三个通道进行图像分析,并得到各个颜色通道的全局纹理特征。局部纹理特征通常是在空域或者变换域通过对图像局部窗口内的像素进行编码并量化而得到的特征。典型地,Ojala 等人^[14]最早提出了局部二值模式 LBP,并用于提取局部纹理特征。Tan 等人^[15]基于局部三

值模式 LTP 提取局部纹理特征,并用于人脸识别。Bu 等人^[16]使用旋转不变的均匀局部二值模式 RULBP 特征用于纹理图像的检索。Wang 等人^[17]提出局部八方向模式 LEDP,并用于纹理图像的检索。此外,Pham 等人^[18]基于局部极值描述符 LED,以较低的特征维度有效地实现了纹理图像的检索。

相比较而言,基于小波类变换域统计建模的全局特征在彩色纹理图像检索领域得到了广泛的应用。小波类变换分为实小波类变换和复小波类变换。与利用实小波类子带系数直接进行建模特征提取的方式不同,复小波类变换域提取的特征包括对子带的幅值与相位信息的描述。业已证明,幅值特征与相位特征反映了图像的能量与结构信息,其中相位特征尤为重要,它蕴含了图像的重要结构与方向信息^[19-20]。然而,在上述纹理特征提取方法中,或有纹理特征提取不全面、不充分,或有统计模型选择不合适、不恰当,或者缺乏有效地利用全局的和局部的相位特征等不足。特别地,融合不同类型的互补相位特征有利于提高图像间的区分能力和图像检索的性能。

近年来,在基于深度学习的实例图像检索方面取得了极大的进展。Gordo 等人^[21]通过分析深层图像检索方法效果不佳的原因提出一种具有抗噪声性能的实例级图像检索方法。Filip 等人^[22]基于卷积神经网络提出一种改进池化层来对网络结构进行微调的检索方法。Zhang 等人^[23]将深度学习与哈希技术相结合,一定程度上提高了检索效率。在基于深度学习的彩色纹理图像检索方面,Bhunia 等人^[24]结合哈希方法用于检索,但获得的结果不及传统方法。究其原因,目前主要是因缺乏大型公认的数据集,从而造成基于深度学习的纹理图像检索研究较少,且已有研究的检索性能不理想。

针对上述问题,该文提出了一种基于全局和局部相位特征互补融合、并结合高效的颜色特征的彩色纹理图像检索方法。首先,在 HSV 颜色空间,以符合人类色彩视觉特点的非均匀量化方式,分别对三个颜色通道进行量化,并构建联合直方图用作图像的颜色特征;同时,针对 V 通道的图像数据,利用 Gabor 复变换域相对相位子带系数建模的方式提取全局纹理相位特征。然后,使用局部引力角模式 PLGFA 描述符,在 RGB 空间提取灰度图像的相位信息作为局部纹理相位特征。最后,通过采用凸线性优化加权方式,将各特征对应的相似性测度有机组合成一个总的相似性测度,进而完成检索任务。

彩色纹理图像检索实验在 VisTex (full/small)、STex 和 ALOT 等四个通用的数据库中进行,实验结果验证了该方法的有效性和可行性。

1 相关工作

1.1 颜色空间

在基于全局和局部相位特征融合的彩色纹理图像检索系统中,选择一个合适的颜色空间并提取有效的特征,对系统性能的影响至关重要。通常,在图像数据存储与表达方面,RGB 颜色空间是一个较好的选择。另一方面,相关的研究成果表明,HSV 颜色空间更符合人类视觉的色彩感知,并且在此空间中所提取的颜色特征用于图像检索系统也更为有效^[25]。在 HSV 颜色空间中,H(Hue)分量表示一幅图像的色调,其使用角度度量,取值范围为 $0^\circ \sim 360^\circ$ 。S(Saturation)分量用来表示一幅图像的色饱和度,其使用百分比度量,取值范围为 $0\% \sim 100\%$ 。它反映了图像颜色接近于光谱色的程度,其值越大,颜色越饱和。V(Value)分量用来表示一幅图像的亮度,其取值范围为 $0 \sim 1$ 。

该文将 RGB 颜色空间与 HSV 颜色空间同时用于彩色纹理图像检索。其中,在 HSV 空间,除了提取图像的颜色特征之外,图像的 V 分量数据还用于全局纹理相位特征的提取;而由 RGB 空间生成的灰度图像则根据 PLGFA 描述符提取局部纹理相位特征。

1.2 Gabor 滤波器

二维 Gabor 滤波器是一组复正弦平面波调制的高斯核函数^[26],其定义如式(1)~式(3)所示:

$$G(x, y) = \frac{f^2}{\pi\gamma\eta} \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp(j2\pi f x' + \varphi) \quad (1)$$

$$x' = x \cos \theta + y \sin \theta \quad (2)$$

$$y' = -x \sin \theta + y \cos \theta \quad (3)$$

式中, f 为正弦波的频率; θ 为法线相对 Gabor 函数平行条纹的方向; φ 为相位偏移; σ 为高斯包络的标准偏差; γ 为空间纵横比,用来决定 Gabor 函数形状的椭圆度; η 为反映高斯核锐度的参数。通常,二维 Gabor 滤波器在空域和频域中同时具有最优的局域化特性。

1.3 PLGFA

局部引力角模式(PLGFA)是根据牛顿万有引力定律衍生的一种描述符^[2]。万有引力定律是指在已知宇宙中,任何两个星体之间都存在相互吸引的力,引力大小与它们质量的乘积成正比、与它们距离的平方成反比。其中,这种引力是两个成对出现且大小相同的、沿两个星体质心连线指向对方的力。对于两个星体之间的引力及其分力如图 1(a) 所示。

直观地,在空域每幅灰度图像的像素类似于星体,像素的灰度值相当于星体的质量,像素与其局域相邻像素之间的相关性关系也类似于星体之间的引力关系,即它们之间的距离越小,相关性越强。因此,对于一幅灰度图像,在任一 3×3 大小的局域内,中心像素

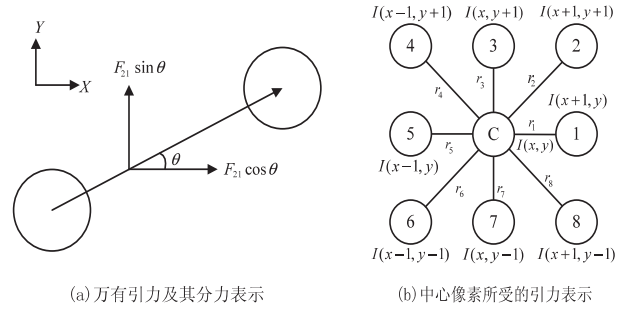


图 1 典型的二维 Gabor 滤波器的基图像和引力表示示意图

c 所受的引力可以由图 1(b) 表示。通常,中心像素的引力可以用其所受合力的大小和方向角(相位角)两个分量进行描述。其中,结合照度-反射率模型(Illumination-Reflectance Model, IRM)^[2],中心像素引力的方向角 α 可以表示为:

$$\alpha = \arctan\left(\frac{F_y}{F_x}\right) = \arctan\left[\frac{\sum_{i=1}^8 \left(G \times \frac{I_c \times I_i}{r_{ci}^2} \times \sin \theta_{ci}\right)}{\sum_{i=1}^8 \left(G \times \frac{I_c \times I_i}{r_{ci}^2} \times \cos \theta_{ci}\right)}\right] = \arctan\left[\frac{\sum_{i=1}^8 (I_i \times \sin \theta_{ci})}{\sum_{i=1}^8 (I_i \times \cos \theta_{ci})}\right] \quad (4)$$

式中, $\alpha \in [0, 2\pi)$; F_x 、 F_y 分别为水平与垂直方向的分力; I_c 为中心像素的灰度值; I_i 为邻域像素的灰度值; r_{ci} 为中心像素与邻域像素之间的欧氏距离; θ_{ci} 为中心像素和邻域像素的引力方向与 x 轴间的夹角; G 为万有引力常量。相对应地,对于每幅图像中的任一像素,均可以由式(4)计算出其局部引力方向角。为了构造 PLGFA,首先对于一个 3×3 大小的邻域,计算邻域像素方向角 $\alpha_i (i = 1, 2, \dots, 8)$ 与中心像素方向角 α_c 差值的绝对值向量 \mathbf{ang} 。即:

$$\mathbf{ang} = [|\alpha_1 - \alpha_c|, |\alpha_2 - \alpha_c|, \dots, |\alpha_8 - \alpha_c|] \quad (5)$$

然后,根据阈值 T ,对此绝对值向量进行二进制编码,并最终由式(6)计算得到中心像素的 PLGFA 值。即:

$$\text{PLGFA} = \sum_{i=1}^K 2^{i-1} \times R(\mathbf{ang}_i) \quad (6)$$

$$R(a) = \begin{cases} 1, & a \leq T \\ 0, & \text{otherwise} \end{cases}$$

式中, K 为 3×3 邻域内邻域像素的个数; \mathbf{ang}_i 为 \mathbf{ang} 向量的第 i 个分量; 阈值 T 由实验确定。相应地,对于每幅图像中的任一像素,均可以由式(4)~式(6)计算出其 PLGFA 值,从而得到对应图像的 PLGFA 谱图。

2 提出的检索系统框架

通常,纹理图像具有丰富的方向性,而相位特征可以充分地描述图像的几何结构方向信息。因此,该文旨在通过有效地融合互补的全局和局部相位特征、并

结合高效的颜色特征,研究一种新的彩色纹理图像检索系统。提出的彩色纹理图像检索系统框架如图2所示,具体组成将由下文给出。

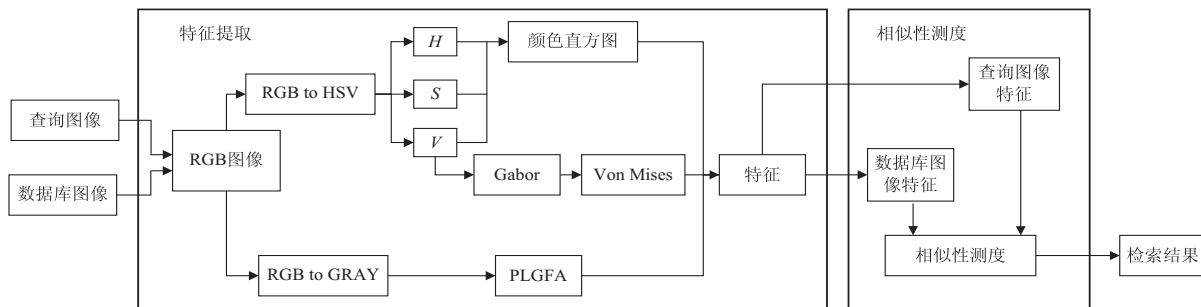


图2 彩色纹理图像检索系统框架

2.1 颜色特征

近年来的众多研究表明,HSV颜色空间相比于其他颜色空间更符合人类的视觉系统。因此,该文只在HSV颜色空间中提取颜色特征^[25]。首先,将彩色纹理图像从RGB颜色空间转换到HSV颜色空间;然后,分别将H、S、V分量量化为8、3、3个级别;最后,通过组合三个分量的量化数据直方图构成颜色特征。三个分量的具体量化方式如式(7)~式(9)所示。即:

$$H = \begin{cases} 7, H \in [0, 14) \\ 6, H \in [14, 28) \\ 5, H \in [28, 42) \\ 4, H \in [42, 56) \\ 3, H \in [56, 72) \\ 2, H \in [72, 198) \\ 1, H \in [198, 324) \\ 0, \text{otherwise} \end{cases} \quad (7)$$

$$S = \begin{cases} 2, S \in (0.5, 0.8] \\ 1, S \in (0.15, 0.5] \\ 0, \text{otherwise} \end{cases} \quad (8)$$

$$V = \begin{cases} 2, V \in (0.5, 0.7] \\ 1, V \in (0.3, 0.5] \\ 0, \text{otherwise} \end{cases} \quad (9)$$

对于量化后的数据,按照H、S、V的顺序,利用式(10)的形式构建一维颜色直方图特征向量。

$$\text{VCH} = \{000, \dots, 020, \dots, 022, \dots, 722\} \quad (10)$$

同时,由颜色特征向量的组合过程可知,VCH是一维、72 bin的直方图特征向量。此外,按照从小到大的排列方式,可以将VCH的值映射到0~71,以便于计算和存储。

2.2 全局的纹理相位特征

有关的生物视觉研究表明,Gabor滤波器可以很好地对人眼视觉细胞的感知特性进行建模^[11]。同时,HSV颜色空间是一个适合人类视觉感知的颜色空间,

其V分量数据能够较好地反映人类视觉对图像的灰度感知。因此,为了提取全局的纹理相位特征,该文使用典型的具有3尺度、每尺度下含有8方向的Gabor滤波器组对V分量数据进行变换,共得到24(3*8)个复值子带。对于所有的复值子带,由于每个子带的实部和虚部均可以由广义高斯分布很好地拟合,因此由实部和虚部计算得到的相对相位子带就可以由Von Mises分布进行统计建模^[20]。假定相对相位子带系数为 θ ,并由两参数的Von Mises分布进行统计建模,则其概率密度函数为^[20]:

$$p(\theta | \mu, v) = \frac{1}{2\pi I_0(v)} e^{v \cos(\theta - \mu)} \quad (11)$$

式中, μ 为平均方向参数; v 为集中程度参数; I_0 为第一类零阶修正贝塞尔函数,且 $I_0(v) = \frac{1}{\pi} \int_0^\pi e^{v \cos \theta} d\theta$ 。

文中Von Mises分布的参数采用最大似然方法进行估计,并将其作为全局的纹理相位特征。对于通过估计所得到参数的准确性,通常使用熵差率 R_e 进行评估^[17]。 R_e 的值越小,说明估计得到的参数值越接近于真实的分布参数值、对模型的估计也越准确。为了验证所采用的相对相位子带分布模型的有效性,分别选取了DB1数据库中的第350幅图像和DB4数据库中的第600幅图像进行拟合实验(其中,数据库DB1~DB4的详情参见第3节)。首先,分别对这两幅图像的V分量进行3尺度、每尺度下8方向的Gabor变换;然后,分别选取变换后第2尺度下的第7个相对相位子带(135°方向)系数进行分布建模的拟合实验,实验结果如图3(a)~(d)所示。由图3(b)和(d)中的拟合结果以及对应的熵差率这一评价指标的数值可见,所采用的Von Mises分布模型可以很好地符合Gabor变换域相对相位子带系数的真实分布,即利用此分布模型的参数作为全局纹理相位特征是合适的,它有利于提高该检索系统的性能。

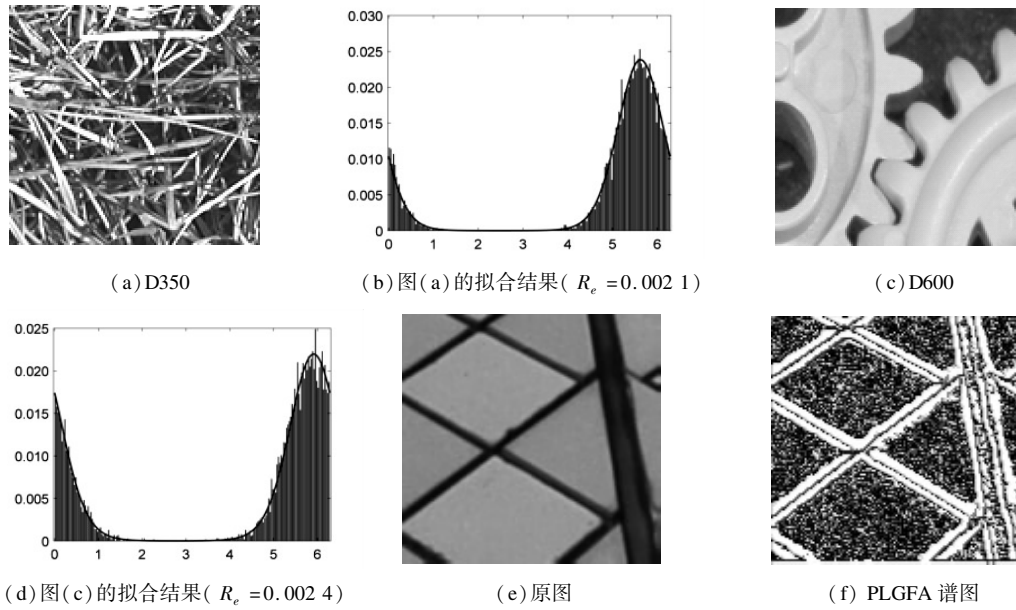


图 3 子带建模拟合效果及 PLGFA 描述符提取图像相位信息效果示意图

2.3 局部的纹理相位特征

对于局部的纹理相位特征,采用 PLGFA 进行描述。特别地,PLGFA 描述符作用于 RGB 颜色空间的灰度图像。因此,需要将 RGB 彩色图像通过式(12)转换成灰度图像。即:

$$I_{\text{gray}} = 0.299R + 0.587G + 0.114B \quad (12)$$

需要说明的是,之所以选择 RGB 颜色空间的灰度图像而不是使用 HSV 颜色空间的 V 分量,其原因是 RGB 空间的灰度图像更符合图像像素之间的自然分布,从而更加契合由万有引力定律导出的 PLGFA 描述符。

基于 PLGFA 的局部纹理相位特征的提取过程如下:首先,根据 RGB 颜色空间的一幅灰度图像 I_{gray} 计算每个像素的局部引力角,进而得到此图像的局部引力角图;然后,经过在各数据库上的测试实验,确定式(6)中的阈值 T ;最后,根据式(6)计算整幅图像的局部引力角模式谱图,并通过直方图量化成一维、256bin 的特征向量。其中,基于 DB1 ~ DB4 四个数据库检索率最优表现,均取 $T = \pi/6$ 。

所采用的 PLGFA 描述符可以充分地提取纹理图像局部的相位特征,其反映的方向信息效果如图 3(f) 例示。由图 3(f) 可见,PLGFA 描述符可以有效地刻画出各个方向的边缘,即保留了纹理图像的方向信息。因此,基于 PLGFA 描述符所提取的局部相位特征有利于检索系统性能的改善。

2.4 相似性测度

对于颜色特征和局部纹理相位特征,由于它们都是直方图特征(参见 2.1 和 2.3 节),因此选取改进版本的 Manhattan 距离进行相似性计算^[27]。式(13)和式(14)分别列出了颜色特征之间以及局部纹理相位

特征之间的相似性测度。即:

$$D_{\text{VCH}}(\text{img}_{\text{db}}, \text{img}_{\text{q}}) = \sum_{i=1}^P \left| \frac{V_{\text{db}}(i) - V_{\text{q}}(i)}{1 + V_{\text{db}}(i) + V_{\text{q}}(i)} \right| \quad (13)$$

$$D_{\text{PLGFA}}(\text{img}_{\text{db}}, \text{img}_{\text{q}}) = \sum_{j=1}^Q \left| \frac{V_{\text{db}}(j) - V_{\text{q}}(j)}{1 + V_{\text{db}}(j) + V_{\text{q}}(j)} \right| \quad (14)$$

式中, img_{db} 和 img_{q} 分别为数据库中的备选图像与查询图像; V_{db} 和 V_{q} 分别为特征数据库中的备选特征向量以及查询图像的特征向量; P 、 Q 为特征向量的长度。

对于全局纹理相位特征,该文采用具有闭式形式的 K-L (Kullback-Leibler) 距离计算两个 Von Mises 分布参数特征向量之间的相似性。式(15)列出了两个 Von Mises 分布间的 K-L 距离^[20]。

$$D_{\text{VonMises}}(\text{img}_{\text{db}}, \text{img}_{\text{q}}) = \log \frac{I_0(v_2)}{I_0(v_1)} + \frac{I_1(v_1) - I_1(-v_1)}{2I_0(v_1)}(v_1 - v_2 \cos(\mu_2 - \mu_1)) \quad (15)$$

式中, img_{db} 和 img_{q} 分别为数据库中的备选图像与查询图像; μ_1 、 v_1 为备选特征数据库 Von Mises 分布模型的两个参数特征; μ_2 、 v_2 为查询图像的 Von Mises 分布模型的两个参数特征; I_1 为第一类一阶修正贝塞尔函数,且 $I_1 = \frac{1}{\pi} \int_0^\pi e^{v \cos \theta} \cos \theta d\theta$ 。

根据提出的三种特征所对应的相似性测度,为了综合评估两幅图像之间的差异性,还需要计算它们所有特征之间总的相似性测度。为此,该文采用凸线性优化组合的方式计算总的相似性测度。即,采用式(16)的形式计算总的相似性测度。

$$D(\text{img}_{\text{db}}, \text{img}_q) = aD_{\text{VCH}} + bD_{\text{VonMises}} + cD_{\text{PLGFA}} \quad (16)$$

式中, $D(\text{img}_{\text{db}}, \text{img}_q)$ 表示数据库中的备选图像与查询图像之间的总相似性测度; a 、 b 、 c 分别为三种不同特征对应的相似性测度在总相似性测度中的权重, 且满足 $a + b + c = 1$ 。其中它们的取值范围均在 0 到 1 之间。该文以 0.01 步长遍历的方式调整三个权重的大小以获得平均检索率的最优值。具体地, 利用 Matlab 软件, 根据提出的方法在 DB1 ~ DB4 四个数据库上分别进行检索实验, 得到的最优权重结果分别为:

DB1: $a = 0.39$, $b = 0.37$, $c = 0.24$;

DB2: $a = 0.25$, $b = 0.63$, $c = 0.12$;

DB3: $a = 0.37$, $b = 0.50$, $c = 0.13$;

DB4: $a = 0.34$, $b = 0.54$, $c = 0.12$ 。

2.5 检索算法

输入: 查询图像

输出: 检索结果

Step1. 选择查询图像 img_q , 将其与数据库备选图像 img_{db} 分别转换到 HSV 颜色空间, 同时获得 RGB 颜色空间的灰度图像。

Step2. 量化 H 、 S 、 V 三个颜色分量, 根据式(7) ~ (10) 构建颜色直方图, 得到图像的一维颜色特征向量。

Step3. 使用 Gabor 滤波器组对 V 分量进行变换, 由式(11) 计算各相对相位子带的 Von Mises 分布参数, 作为全局的纹理相位特征。

Step4. 使用式(5) 计算灰度图像的局部引力角图, 使用式(6) 并结合阈值 T 计算局部引力角模式谱图, 并通过直方图量化成一维、256 bin 的局部纹理相位特征。

Step5. 将 Step2 ~ 4 得到的特征融合成图像的总特征向量, 使用式(13) ~ (15) 分别计算不同特征相应的相似性测度, 并使用式(16) 计算查询图像与数据库中备选图像之间总的相似性测度。

Step6. 根据数据库, 检索出最相似的 M 个图像(文中取 $M = 16$)。

3 实验结果与讨论

为了评估所提出的彩色纹理图像检索系统的有效性, 该文采用四个被广泛使用的彩色纹理图像数据库进行实验。它们分别为 VisTex (full)、VisTex (small)、STex 和 ALOT 数据库。为了描述和使用方便, 对于重新处理后的这些数据库, 分别用 DB1、DB2、DB3 和 DB4 对应表示。

VisTex (full) 数据库由 167 幅、每幅大小为 $512 * 512$ 像素的 RGB 图像构成。它包含了麻省理工学院

视觉纹理数据库 (VisTex) 中的所有图像, 其中包括植物、纺织品等物体的影像。VisTex (small) 数据库是 VisTex (full) 数据库的一个子集, 它由 40 幅图像组成。典型地, 将两个数据库中的每一幅图像无重叠地分割成 16 幅、每幅大小为 $128 * 128$ 像素的 RGB 子图像, 并由此构成包含 2 672 ($167 * 16$) 幅与 640 ($40 * 16$) 幅彩色纹理图像的实验数据库 DB1 和 DB2。

STex 是一个具有 476 幅、每幅大小为 $512 * 512$ 像素的 RGB 图像的数据库, 其中包含食物、石头、木头等物体的影像。DB3 由 STex 数据库构成的方式与上述数据库的相同, 它是一个包含 7 616 ($476 * 16$) 幅、每幅大小为 $128 * 128$ 像素的彩色纹理图像实验数据库。

ALOT 为包含了 250 幅、每幅大小为 $768 * 512$ 像素的 RGB 图像的数据库, 它含有结晶体、编织品、硬币等物体的影像。对于 ALOT 数据库中的每幅图像, 首先采用居中的方法裁剪出 $512 * 512$ 像素大小的 RGB 图像; 然后再对其进行无重叠的 16 等分; 最终得到包含 4 000 ($250 * 16$) 幅、每幅大小为 $128 * 128$ 像素的 RGB 子图像的数据库 DB4。图 4 是四个数据库中部分图像的示例。

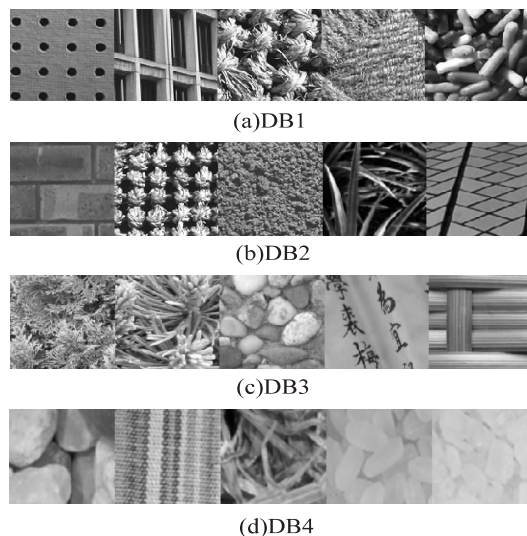


图4 四个数据库中部分图像的示例

为了比较该文所提出的方法与其他相关的最新方法间的检索性能, 将采用平均检索率 (Average Retrieval Rate, ARR) 作为评价指标。ARR 常以百分比的形式表示, 其计算方式如式(17) 所示。

$$\text{ARR} = \frac{1}{K} \sum_{m=1}^K \frac{N_{16}(I_m)}{16} \quad (17)$$

式中, K 表示数据库中图像的数量; $N_{16}(I_m)$ 表示在第 m 次检索的前 16 个结果里, 与查询图像 I_m 同属一类图像的数量。

3.1 消融实验

为了验证相位特征信息对于改善彩色纹理图像检索性能的重要性, 采用四种实验方案进行了最优 ARR

的对比。由于在第 2.2 节提出的检索系统中所使用的 Gabor 滤波器产生的子带较多,容易在提取的特征中造成较多的冗余,因此在后续实验中只选取了部分代表性子带进行特征提取。这些子带分别为每一个分解尺度的 0° 、 45° 、 90° 、 135° 方向子带。这样,仅有 $12(3 * 8/2)$ 个子带参与检索,从而一方面减少了特征冗余,另一方面降低了特征提取的时间开销。在消融实验中,首先选用了两种幅值模型和两种相位模型,分别是:(1)Gabor 变换域的幅值模型(分解 3 层、每层 4 个幅值子带、特征为幅值 Gamma 分布参数、相似性测度为 K-L 距离);(2)局部引力幅值模式模型(特征是局部引力幅值模式直方图、相似性测度是改进版本的 Manhattan 距离);(3)Gabor 变换域的相对相位模型(分解 3 层、每层 4 个相对相位子带、特征为相对相位 Von Mises 分布参数、相似性测度为 K-L 距离);(4)局部引力角模式模型(特征是局部引力角模式直方图、相似性测度是改进版本的 Manhattan 距离)。然后,为了简明,对于这 4 种模型采取两两适当结合的方法分别在 DB2 数据库上进行检索实验,实验结果如表 1 所示。

表 1 四种组合方案的 ARR 比较

方法	ARR/%
(1)+(2)	91.03
(1)+(3)	91.39
(2)+(4)	92.68
(3)+(4)	93.29

通过对比表 1 的四组实验数据可见,基于相位特征的方法组合较其他三种组合的 ARR 值分别高出

2.26%、1.90% 和 0.61%,因此,此组合方案在四种方案中是最好的,这也是该文所提出方法的依据。在后续针对四个数据库的检索实验中,均采用由两种相位模型组成的检索方案。

3.2 主观评价

彩色纹理图像检索系统的性能既可以根据客观指标进行定量比较评价,又可以从视觉效果来进行定性的主观评估。首先,根据所提出的方法(Proposed Method, PM)在四个数据库上进行检索实验,并选取部分查询图像的检索结果进行定性分析。

图 5(a)~(d)为 PM 方法在四个数据库上的部分检索结果,图中实线框出的是与查询图像相同的图像,虚线框出的是与查询图像不同类的图像,即查询到的错误图像。图 5(a)和图 5(b)展示了 PM 方法在 DB1、DB2 数据库中针对第 577、337 幅查询子图像的较好检索结果。说明 PM 方法可以在数据库中准确地找到目标图像。同时,通过观察图 5(a)的错误图像,发现它们与其他检索到的同类图像间具有非常相似的视觉效果,同时颜色也比较接近,这说明 PM 方法在检索彩色纹理图像时具有较强的语义信息捕捉能力。

图 5(c)和图 5(d)例示了 PM 方法在 DB3、DB4 数据库中针对第 141、114 幅查询子图像的检索结果。需要说明的是,图 5(c)和图 5(d)中选取的是检索效果最差的两幅查询图像。但从检索结果可见,在检索出的前 5 幅图像中,与查询图像同一类的比例仍然是较高的。这说明 PM 方法即使在一些数据库上的 ARR 较差,但是仍然可以有效地检索出最相似的若干目标图像。

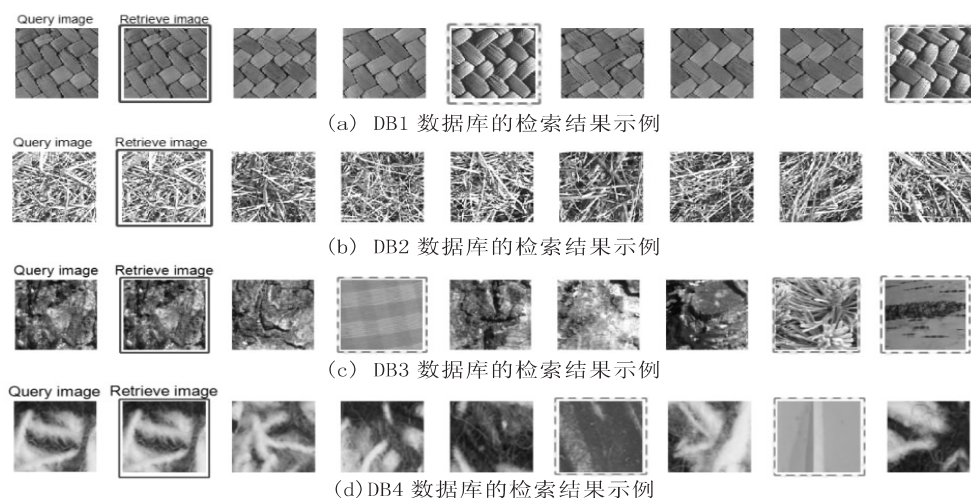


图 5 四个数据库的检索结果示例

3.3 不同方法的检索性能比较

为了进一步说明所提出的 PM 方法的有效性,针对上述四个数据库 DB1~DB4,将其与现有的、密切相关的一些最新方法进行检索性能的对比,并以 ARR

作为客观评价指标。其中,将 PM 方法分别与基于手工算法中的“Gabor - MDCM/RD^[1]”(RD 为 Riemannian 距离)、“MTH/L1^[3]”、“KBH/CSD^[3]”(CSD 为 Chi-Square 距离)、“LEPVP + CH/L1^[4]”、

“LECoP + HSV/D1^[5]”、“DSCoP + MCH/D1^[6]”、“MCLTP/D1^[7]”、“DDBTC/MCD^[8]”(MCD为改进的Canberra距离)、“CMGW/KLD^[10]”、“GCW/ML^[10]”等方法,以及基于深度学习的“VGG19/ED^[1]”方法(ED为Euclidean距离)进行比较,实验结果如图6所示。

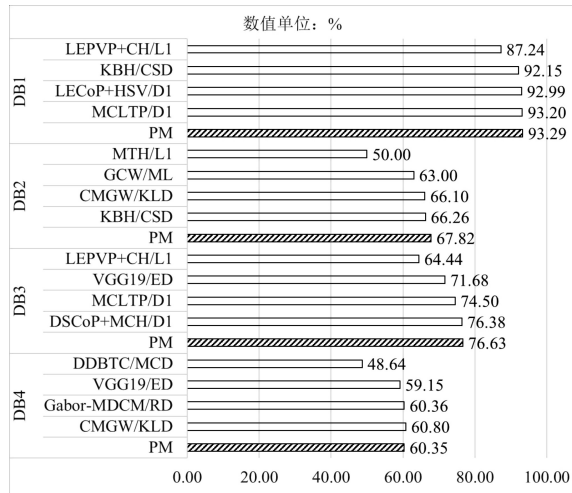


图6 PM方法与现有最新方法在4个数据库上的ARR性能比较(%)

由图6的数据可见,通过与现有最新方法的ARR值相对比,PM方法在DB1、DB2与DB3数据库上具有相当的优势;而在DB4数据库上,尽管PM方法的ARR结果稍逊于诸如Gabor-MDCM和CMGW等同样使用Gabor变换的方法,但是特征提取仅采用了3个尺度、每尺度下8个方向中一半的子带,即12个子带(详见3.1节),而上述两种方法分别使用了40个(5个尺度、每尺度下8个方向)和32个(4个尺度、每尺度下8个方向)子带用于特征提取。因此,相对于较少的ARR值提升,PM方法所节省的计算量更具有竞争性。

3.4 检索效率

为了验证和评估PM方法的检索效率,选取了部分现有的最新方法进行特征长度和检索时间消耗的比较。其中,实验是在搭配32G RAM的具有英特尔酷睿I7处理器的系统平台以及Matlab 2020a软件平台上完成。首先,对比了部分方法的特征长度,如表2所示。

表2 PM方法与现有最新方法特征维度的对比

方法	维度
LEPVP+CH/L1	248
DSCoP+MCH/D1	348
LECoP+HSV/D1	294
PM	372

由表2中的数据可见,所提出的方法在特征向量

的维度这一指标上并不占有优势。但是,考虑到所采用的特征中有328(72+256)个维度为直方图特征,而这些特征的计算耗时并不多。相反,尽管其他方法的特征维度较低,但是它们的特征提取计算过程远比文中的繁复,相应地耗时就较高。因此,为了全面地衡量检索效率,需要对各种方法对应系统的各部分检索耗时进行进一步的对比,实验结果如表3所示。

表3 PM方法与现有最新方法的检索耗时对比

(time/s)

方法	CMGW/ KLD	LECoP+ HSV/D1	KBH/ CSD	PM
FE	2.664 0	0.241 4	0.144 9	0.062 3
SM	32.986 0	0.54	0.148 1	0.038 3
TE+SM	35.650 0	0.781 4	0.293 0	0.100 6

由图6和表3中的数据可见,相比于CMGW方法,PM方法在ARR值上提高了1.72%,同时在时间消耗上也有较大的优势;而同其他两种方法相比较,PM方法在这两个性能指标方面也有一定的优势。

综上,通过在ARR、特征维度以及检索耗时等三个方面与现有的最新方法进行了检索性能的比较,实验结果表明,文中方法取得了较好的检索性能。

4 结束语

为了解决目前彩色纹理图像检索系统中存在的颜色特征提取不充分、纹理特征提取与匹配的有效性亟待提高等问题,提出了一种基于全局和局部相位特征融合的彩色纹理图像检索方法。首次将PLGFA描述符有效地用于彩色纹理图像的检索。在该方法中,首先通过将RGB空间的彩色纹理图像转换到HSV颜色空间、非均匀量化其三个颜色分量、计算组合颜色直方图而得到颜色特征。然后,由Gabor滤波器组将V分量数据复变换为相对相位子带系数,并通过对其进行Von Mises分布建模而得到全局纹理相位特征;同时,针对RGB空间的灰度纹理图像,由其PLGFA谱图的直方图得到图像的局部纹理相位特征。最后,选取不同的距离计算不同特征对应的相似性测度,并将其凸线性优化组合成总相似性测度,从而完成检索任务。检索性能评价实验在VisTex(full/small)、STex和ALOT等四个通用的彩色纹理图像数据库上进行。实验结果表明,该方法取得了优良的检索效果,且和现有的最新方法相比较,平均检索率有较好的提升、检索效率较高,从而验证了该方法的有效性和可行性。

参考文献:

- [1] LI Chaorong, HUANG Yuanyuan, YANG Xingchun, et al. Marginal distribution covariance model in the multiple wave-

- let domain for texture representation [J]. Pattern Recognition, 2019, 92: 246–257.
- [2] BHATTACHARJEE D, ROY H. Pattern of local gravitational force (PLGF): a novel local image descriptor [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(2): 595–607.
 - [3] BAI C, ZHANG J, LIU Z, et al. K-means based histogram using multiresolution feature vectors for color texture database retrieval [J]. Multimedia Tools and Applications, 2015, 74(4): 1469–1488.
 - [4] DEY M, RAMAN B, VERMA M. A novel colour- and texture-based image retrieval technique using multi-resolution local extrema peak valley pattern and RGB colour histogram [J]. Pattern Analysis & Applications, 2016, 19(4): 1159–1179.
 - [5] VERMA M, RAMAN B, MURALA S. Local extrema co-occurrence pattern for color and texture image retrieval [J]. Neurocomputing, 2015, 165: 255–269.
 - [6] BHUNIA A K, BHATTACHARYA A, BANERJEE P, et al. A novel feature descriptor for image retrieval by combining modified color histogram and diagonally symmetric co-occurrence texture pattern [J]. Pattern Analysis and Applications, 2020, 23: 703–723.
 - [7] AGARWAL M, SINGHAL A, LALL B. Multi-channel local ternary pattern for content-based image retrieval [J]. Pattern Analysis and Applications, 2019, 22(4): 1585–1596.
 - [8] GUO J M, PRASETHO H, WANG N J. Effective image retrieval system using dot-diffused block truncation coding features [J]. IEEE Transactions on Multimedia, 2015, 17(9): 1576–1590.
 - [9] VINAYAK V, JINDAL S. CBIR System using color moment and color auto-correlogram with block truncation coding [J]. International Journal of Computer Applications, 2017, 161(9): 1–7.
 - [10] LI C, HUANG Y, ZHU L. Color texture image retrieval based on Gaussian copula models of Gabor wavelets [J]. Pattern Recognition, 2017, 64: 118–129.
 - [11] TAI S L. Image representation using 2D Gabor wavelets [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1996, 18(10): 959–971.
 - [12] 曲怀敬, 王恒斌, 徐佳, 等. 基于 DTCWT 域统计特征融合的纹理图像检索 [J]. 山东建筑大学学报, 2020, 35(3): 28–35.
 - [13] ETEMAD S, AMIRMAZLAGHANI M. Color texture image retrieval based on Copula multivariate modeling in the Shearlet domain [J]. Engineering Applications of Artificial Intelligence, 2021, 102(23): 104256.
 - [14] OJALA T, PIETIKINEN M, MENP T. Gray scale and rotation invariant texture classification with local binary patterns [C]//European conference on computer vision. Dublin: Springer, 2002: 971–987.
 - [15] TAN X, TRIGGS B. Enhanced local texture feature sets for face recognition under difficult lighting conditions [J]. IEEE Transactions on Image Processing, 2010, 19(6): 1635–1650.
 - [16] BU H H, KIM N C, LEE B H, et al. Content-based image retrieval using texture features extracted from local energy and local correlation of Gabor transformed images [J]. Journal of Information Processing Systems, 2017, 13(5): 1372–1381.
 - [17] WANG H B, QU H J, XU J, et al. Combining statistical features and local pattern features for texture image retrieval [J]. IEEE Access, 2020, 8: 222611–222624.
 - [18] PHAM M T, MERCIER G, BOMBRUN L, et al. Texture and color-based image retrieval using the local extrema features and riemannian distance [EB/OL]. (2017-03-03) [2021-07-30]. <https://arxiv.org/abs/1611.02102>.
 - [19] 刁麓弘. 图像处理中的相位信息研究与应用 [D]. 北京: 中国科学院计算技术研究所, 2007.
 - [20] VO A, ORAINTARA S. A study of relative phase in complex wavelet domain: Property, statistics and applications in texture image retrieval and segmentation [J]. Signal Processing Image Communication, 2010, 25(1): 28–46.
 - [21] GORDO A, ALMAZÁN J, REVAUD J, et al. End-to-end learning of deep visual representations for image retrieval [J]. International Journal of Computer Vision, 2017, 124: 237–254.
 - [22] FILIP R, GIORGOS T, ONDREJ C. Fine-tuning CNN image retrieval with no human annotation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 41: 1655–1668.
 - [23] ZHANG Z, ZOU Q, LIN Y, et al. Improved deep hashing with soft pairwise similarity for multi-label image retrieval [J]. IEEE Transactions on Multimedia, 2020, 22(2): 540–553.
 - [24] BHUNIA A K, KISHORE P, MUKHERJEE P, et al. Texture synthesis guided deep hashing for texture image retrieval [C]//2019 IEEE winter conference on applications of computer vision (WACV). Waikoloa Village: IEEE, 2019: 609–618.
 - [25] SINGH S, BATRA S. An efficient bi-layer content based image retrieval system [J]. Multimedia Tools and Applications, 2020, 79(5): 17731–17759.
 - [26] HAGHIGHAT M, ZONOUZ S, ABDEL-MOTTALEB M. CloudID: trustworthy cloud-based and cross-enterprise biometric identification [J]. Expert Systems with Applications, 2015, 42(21): 7905–7916.
 - [27] VERMA M, RAMAN B. Local neighborhood difference pattern: a new feature descriptor for natural and texture image retrieval [J]. Multimedia Tools and Applications, 2017, 77(10): 11843–11866.