

一种有效的多尺度纹理识别方法^{*}

孙俊, 何发智, 陈晓, 袁志勇

(武汉大学 计算机学院, 湖北 武汉 430072)

通讯作者: 何发智, E-mail: fzhe@whu.edu.cn

摘要: 局部二值模式(local binary patterns, 简称 LBP)是一种有效的纹理描述算子, 具有算法复杂度低、消耗内存、原理清晰的优点. Damper-Shafter 的证据理论满足比贝叶斯概率论更弱的条件, 具有直接表达“不确定”和“不知道”的能力. 提出了一种结合二者优势的纹理识别方法. 该方法首先计算图像金字塔, 并利用多尺度 LBP 去测量两个纹理图像之间的相似度; 然后, 通过融合每个测试样本的信息来组合测试纹理与每个训练样本相似性度量; 最后, 识别结果由不同纹理之间的最大证据类决定. 实验结果表明, 该方法对给定的图像数据集 1 和数据集 2 分别取得了 96.43% 和 91.67% 的正确率, 优于最初基于 LBP 的纹理识别方法.

关键词: 纹理识别; LBP; Damper-Shafter 证据理论; 信息融合

中文引用格式: 孙俊, 何发智, 陈晓, 袁志勇. 一种有效的多尺度纹理识别方法. 软件学报, 2014, 25(Suppl. (2)): 278-289. <http://www.jos.org.cn/1000-9825/14045.htm>

英文引用格式: Sun J, He FZ, Chen X, Yuan ZY. Efficient multi-scale texture recognition algorithm. Ruan Jian Xue Bao/ Journal of Software, 2014, 25(Suppl. (2)): 278-289 (in Chinese). <http://www.jos.org.cn/1000-9825/14045.htm>

Efficient Multi-Scale Texture Recognition Algorithm

SUN Jun, HE Fa-Zhi, CHEN Xiao, YUAN Zhi-Yong

(School of Computer Science, Wuhan University, Wuhan 430072, China)

Corresponding author: HE Fa-Zhi, E-mail: fzhe@whu.edu.cn

Abstract: As an effective texture description operator, local binary patterns (LBP) has the advantages of low computation complexity, low memory consumption and clear principle. Damper-Shafter evidence theory satisfies the conditions weaker than Bayesian probability theory and can directly express states of “uncertain” and “don’t know”. To exploit the advantages of above two concepts, a new texture recognition method is proposed. Firstly, the approach computes image pyramid and uses the distributions of multi-scale LBP to measure the similarity between two texture images. Secondly, the method combines the similarity measurement between the test texture and each training sample to combine the information given by each training sample. Finally, the recognition result is determined by the maximum evidence among different texture classes. Experimental results show that the proposed method achieves a correction rate of 96.43%, and 91.67%, for data set 1 and data set 2, respectively, outperforming the original LBP based texture recognition algorithm.

Key words: texture recognition; LBP; Damper-Shafter evidence theory; information fusion

人类可以通过纹理获取大量的视觉场景的信息. 二维纹理分析在现实生活中有许多重要的应用, 例如工业表面检测、遥感和生物医学图像分析. 然而, 现实中对于纹理识别和分类依然存在一定的困难, 其主要原因是, 在现实世界中, 纹理通常会由于位置方向、缩放或光照等因素发生变化而变化. 另一个导致纹理识别困难的原因是纹理种类繁多, 众多的种类可能会导致训练样本在特征空间的分散, 从而导致误判.

目前, 研究者已提出了大量的纹理识别方法. 灰度共生矩阵(gray-level co-occurrence matrix, 简称 GLCM) 方法是基于像素分布强度的统计研究^[1]. 通过考虑像素对在特定的空间中彼此间的关系, 从而形成共生矩阵的原始数据像素对所获得的二阶统计量. 其中纹理特征的统计量定义有 14 种, 但仅有 5 个特征统计量经常使用, 分别

* 基金项目: 国家自然科学基金(61472289); 国家重点基础研究发展计划(973)(2011CB707904)

收稿时间: 2014-05-09; 定稿时间: 2014-08-19

是能量、熵、相关性、同质性和惯性^[2]。

在 Connors^[3]的文章中,一个两维的 Gabor 滤波器组被用来分析纹理.滤波器组由一个正弦平面波的某一频率和方向组成,通过一个二维高斯包络调制.用于该算法的滤波器导致频率接近均匀的覆盖,并且构成了小波变换的近似计算.多通道滤波的纹理分析方法的优点是滤波后的图像统计简单,可以作为纹理特征.

局部二值模式(local binary pattern,简称 LBP)首次作为纹理算子比较研究的一部分提出来^[4],它是一种用来描述图像局部纹理特征的算子.LBP 构造了一种衡量一个像素点和它周围像素点的对比关系.

最初的 LBP 算子的局限在于它的空间支持区域较小.在局部的 3×3 邻域内,计算功能无法捕捉大尺度结构.同时,算子针对纹理的局部变化,诸如视点变化及光流方向等因素而导致算法的不鲁棒性.针对这种局限性,M. Pietikäinen^[5]提出了一种改进算法,通过一种简单的扩大空间支持区域的方法,通过组合多个 LBP 算子周围不同的区域.然而,这种方法存在的高维运算影响了算法的运行速度.同时,在扩展 LBP 算子的基础上分析其结构和发生概率相结合的 non-uniform 局部模式在文献[6]中得到实现.新的算子充分利用包含在 non-uniform 局部模式的纹理信息——这是由经典的 LBP 算子所丢弃的,新的算子也具有更强大的抗噪能力.

除了特征提取以外,监督模式分类是模式识别的另一个讨论的话题. K 近邻(K -nearest neighbor,简称 KNN)分类器也是一种常用的纹理分类方法.此外,支持向量机(support vector machine,简称 SVM)方法已经在实践中得到广泛的应用.Vapnik^[7]及 Li 等人^[8]分别对 SVM,KNN 在纹理分类方面更广义的分类应用进行了研究.然而,随着纹理数据库中的图像数量不断增加,由于分类种类的变化,会造成分类精度下降.

近年来,动态纹理的研究也逐渐成为热点,Derpanis 和 Wildes^[9,10]提出了动态纹理的表示及识别的方法.这种方法基于面向结构的时空分布.

本文提出了一种基于 LBP 和 Dempster-Shafer 证据理论(简称 D-S 证据理论)的纹理识别方法.本方法首先计算输入图像的图像金字塔,然后将 LBP 算子施加在每个尺度的图像金字塔上.然后,每个尺度图像金字塔将产生一个特征向量.通过 D-S 证据理论将多尺度信息在每个尺度上通过相似性融合进行整合.特别是,测试图像和目标样本之间的相似度是通过融合测试图像和每个尺度的样品之间的相似度进行计算.

1 纹理识别方法

1.1 局部二值模式

基本局部二值模式算子是基于一个纹理有局部两个相辅相成的方面,即模式及其强度的假设.原始的 LBP 算子定义为在 3×3 的窗口内,以窗口中心像素为阈值,将相邻的 8 个像素的灰度值与其进行比较.若周围像素值大于中心像素值,则该像素点的位置被标记为 1,否则为 0.这样, 3×3 邻域内的 8 个点可产生 8bit 的无符号数,即得到该窗口的 LBP 值,并用这个值来反映该区域的纹理信息.如图 1 所示.

Image block	Thresholded	Weights																											
<table border="1"> <tr><td>60</td><td>10</td><td>20</td></tr> <tr><td>55</td><td>60</td><td>50</td></tr> <tr><td>128</td><td>82</td><td>35</td></tr> </table>	60	10	20	55	60	50	128	82	35	<table border="1"> <tr><td>1</td><td>0</td><td>0</td></tr> <tr><td>0</td><td></td><td>0</td></tr> <tr><td>1</td><td>1</td><td>0</td></tr> </table>	1	0	0	0		0	1	1	0	<table border="1"> <tr><td>1</td><td>2</td><td>4</td></tr> <tr><td>128</td><td></td><td>8</td></tr> <tr><td>64</td><td>32</td><td>16</td></tr> </table>	1	2	4	128		8	64	32	16
60	10	20																											
55	60	50																											
128	82	35																											
1	0	0																											
0		0																											
1	1	0																											
1	2	4																											
128		8																											
64	32	16																											
Pattern = 01100001 LBP = 64+32+1 = 97																													

Fig.1 An example of basic LBP operator

图 1 一个基本 LBP 算子实例

通用 LBP 算子对邻域的大小或采样点的数目没有限制.对于给定的图像 I ,考虑在 P 采样点的等间距圆域内的任意像素点 (x,y) 和半径 R 的圆形邻域内的像素 (x,y) .像素 (x,y) 和采样点周围像素 (x,y) 的灰度级分别记为 g_c 和 g_p .通用 LBP 算子由式(1)导出,像素及其空间关系由式(2)给定.图 2 为 LBP 算子应用于图片处理的实例,图 3 为通用 LBP 算子的一些应用实例.

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (1)$$

$$\begin{cases} g_p = I(x_p, y_p), p = 0, \dots, P-1 \\ x_p = x + R \cos(2\pi p / P) \\ y_p = y - R \sin(2\pi p / P) \end{cases} \quad (2)$$

其中,式(1)中的 S 是符号标志符.

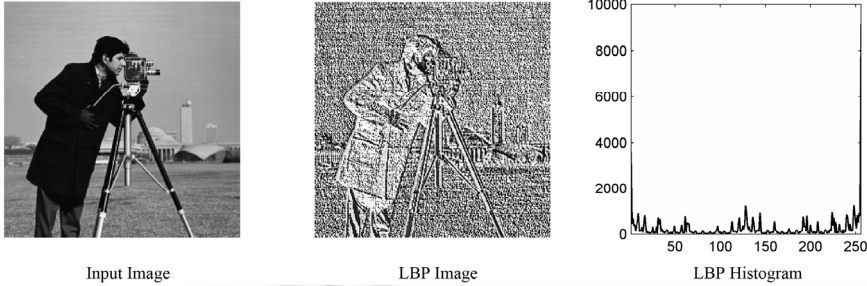


Fig.2 Example of an input image, the corresponding LBP image and histogram

图2 输入图片及相应的 LBP 图像和 LBP 直方图

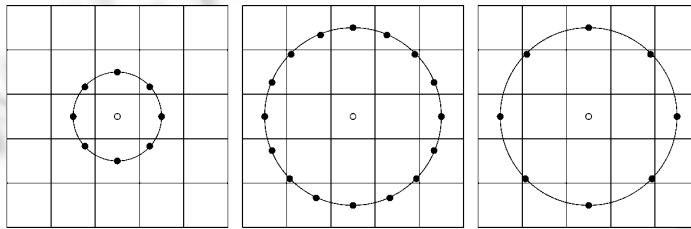


Fig.3 LBP operators circular (8,1), (16,2) and (8,2) neighborhoods

图3 通用 LBP 算子圆形(8,1),(16,2)和(8,2)邻域例子

在 Ojala 等人的另一篇文献^[11]中,基本 LBP 算子通过旋转不变的映射生成旋转不变的 LBP 算子.为了形成旋转不变的编码模式,让有同一编码模式经旋转后产生的编码结果为同一值,即这些旋转结果中的最小值:

$$LBP_{P,R}^r = \min_i ROR(LBP_{P,R}, i) \quad (3)$$

其中, $ROR(x, i)$ 表示位序列 x 循环右旋转的第 i 步.

1.2 Dempster-Shafer 证据理论

D-S 证据理论最初在 Dempster 的文献中通过利用上、下限概率来解决多值映射问题引入^[12].随后,Shafer 完善了这一理论^[13],并且显示出这一理论对于运用信任函数模拟不精确和不确定数据的优势.

在证据理论的背景下,不精确和不确定数据通过信任结构建模,即信任和合理性,每一种都来源于基本信任结构(也称为基本信任分配).这些函数的存在和使用包括识别 Ω 框架的定义,它是由 N 个详尽的和排他的假设构成,即 $\Omega = \{H_1, H_2, \dots, H_N\}$. 所有 2^N 个定义在 Ω 上的命题集 A 的幂集由式(4)给出:

$$2^\Omega = \{\emptyset, H_1, H_2, \dots, \{H_1 \cup H_2\}, \{H_1 \cup H_3\}, \dots, \Omega\} \quad (4)$$

D-S 证据理论提出了用一个框架去量化一些代理的观点,这个给定命题 A 的代理为真.其中, A 既可以是单元素集合 H_n 也可以是分离的假设.被信源 S 设定的基本信任分配 m 定义如下:

$$\begin{cases} m: 2^\Omega \rightarrow [0, 1] \\ \sum_{A \subseteq \Omega} m(A) = 1 \end{cases} \quad (5)$$

其中, $m(A)$ 用于测量信任度的数量. 由于分离假设 (A 作为一个复合假设), A 是证据 $m(A)$ 在缺乏信息的任何子集中不可能重新分配的一部分, 仅仅当新的信息注入, 信任度更精确时, 元素 $A \subseteq \Omega$, 如 $m(A) \neq 0$ 被称为焦元. 信任函数 (Bel) 和似然函数 (Pl) 的定义如式(6)所示:

$$\begin{cases} Bel(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B) \\ Pl(A) = \sum_{A \cap B \neq \emptyset} m(B) \end{cases} \quad (6)$$

其中, $Bel(A)$ 即为命题 A 中全部信任的总和, 而似然函数 $Pl(A)$ 则是用于度量分配给 A 的最大可能信任度. 因此, 信任度和似真度是两个双重的概念, 它们之间可以通过 Möbius 变换相互转换^[14].

证据理论提供了可靠的工具去组合不同信息源提供的知识. 组合的正交规则也被称为 Dempster 规则, 也是第一个被定义在证据理论框架中的组合. 定义: m_1, \dots, m_j , 其中, 信任集合 J 来自 J 个不同的信息源. 则通过 Dempster 组合规则组合的 J 个信息源得到的信任函数定义如下:

$$m_{\oplus} = m_1 \oplus \dots \oplus m_i \oplus \dots \oplus m_j \quad (7)$$

其中, 操作算子 \oplus 表示关联的和可交换的. 对于所有 Ω 中的 A , $m_{\oplus}(A)$ 定义为

$$m_{\oplus}(A) = \frac{m_{\cap}(A)}{1-k}, \quad \forall A \subseteq \Omega, \quad (8)$$

其中, m_{\cap} 和 k 分别由下式给出:

$$m_{\cap}(A) = \sum_{A_1 \cap \dots \cap A_j = A} m_1(A_1) \cdot m_2(A_2) \cdot \dots \cdot m_j(A_j) \quad (9)$$

$$k = \sum_{A_1 \cap \dots \cap A_j = \emptyset} m_1(A_1) \cdot m_2(A_2) \cdot \dots \cdot m_j(A_j) \quad (10)$$

其中, 表达式 k 可以解释为 J 个组合信息源之间的冲突因子, 并且在组合中被当作一个归一化因子. 当它的值较大时 ($k \approx 1$) 组合信息源将导致无意义的非连续^[15,16]. 而且, 当 $k=1$ 时, 数据源将会完全对立, 并且数据整合也将无法实现.

1.3 基于LBP和D-S证据理论的纹理识别方法

最初的 LBP 算子一个非常明显的局限性是它的空间支持区域非常小. 在一个 3×3 的领域内计算的特征不能捕获更大尺度的结构, 而且, 算子对于纹理局部视点或者光流方向变化鲁棒性不强.

Pietikäinen 提出一个直接的方式以扩大空间支持区域. 这种方法通过将 N 个 LBP 算子与变化的 P 和 R 值进行组合. 但是这种直接的方式将带来高维的运算成本. 本文提出的方法通过应用图像金字塔^[17,18]来克服上述缺点. 图像金字塔通过一个预定义的低通滤波器对输入图像进行迭代滤波, 尤其当输入图像 I 成为金字塔的最底层或者第 0 层时. 第 1 层包含图像 I_1 , 它是图像 I 的一个滤波版本. 第 2 层的图像 I_2 是通过图像 I_1 滤波获得. 这个过程通过式(11)来描述:

$$\begin{cases} I_l(i, j) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) I_{l-1}(2i+m, 2j+n) \\ I_0 = I \end{cases} \quad (11)$$

当图像金字塔计算出来后, 将基本的 LBP 算子应用到图像的每一个尺度上. 对于任意输入的图像, 将产生若干个直方图, 如 $\{H_0, H_1, \dots, H_L\}$, 对每个与之匹配尺度的直方图, L 代表图像金字塔的最高层. 两个直方图之间的相似度用直方图相交进行计算得到:

$$S(H_X, H_Y) = \sum_{n=1}^B \min \left(\frac{H_X}{\sum_{n=1}^B H_X}, \frac{H_Y}{\sum_{n=1}^B H_Y} \right) \quad (12)$$

其中, H_X, H_Y 分别表示测试图像和训练图像的直方图, B 是两个直方图中 bin 的数量.

测试图像的判别由全部尺度的基本信任规则进行融合来决定. 特别地, 当考虑一个 N 类纹理图像数据库时, 表示如下:

$$DB = \{C_1, C_2, \dots, C_N\}, C_i = \{I^{i1}, I^{i2}, \dots, I^{iM_i}\} \quad (13)$$

其中, M_i 表示在 C_i 种类型的数据库中的样本数量值.

首先,通过计算每个纹理图像的图像金字塔提取相应图像的特征,然后对图像金字塔应用基本 LBP 算子.在某一尺度 s 上,显示某一测试图像 T 属于类别 C_i 相似程度的基本信任结构由式(14)来确定:

$$\left\{ \begin{array}{l} m_s(T \in C_i) = \frac{\frac{1}{M_i} \sum_{j=1}^{M_i} S(H_T, H_{I_{ij}})}{\sum_{i=1}^N \left[\frac{1}{M_i} \sum_{j=1}^{M_i} S(H_T, H_{I_{ij}}) \right]} \\ m_s(T \in C_i \cup C_j) = 0, \quad i \neq j \end{array} \right. \quad (14)$$

式(14)的分子是测试图像与种类 C_i 中所有训练图像平均相似度,分母是一个归一化常数.式(14)可以理解为:在某一尺度 s 上,一个测试图像属于种类 C_i 的相似度与测试图像 T 同种类 C_i 中所有训练图像的平均相似性是成比例的.尤其当每一种类的训练图像数量是一个常量时,如, $M_1 = M_2 = \dots = M$, 式(12)可以简化成如下形式:

$$\left\{ \begin{array}{l} m_s(T \in C_i) = \frac{\sum_{j=1}^{M_i} S(H_T, H_{I_{ij}})}{\sum_{i=1}^N \sum_{j=1}^{M_i} S(H_T, H_{I_{ij}})} \\ m_s(T \in C_i \cup C_j) = 0, \quad i \neq j \end{array} \right. \quad (15)$$

显然,式(14)和式(15)满足式(4),其中,式(4)是任何基本信任分配的必需条件.在尺度 s 上的基本信任分配 m_s 可以当作一个信息源去判定测试图像 T 属于哪一类.因此,我们可以利用 D-S 证据理论去组合各尺度提供的信息源.即

$$m_{\oplus} = m_1 \oplus m_2 \oplus \dots \oplus m_L \quad (16)$$

最终的判定由设定的分配的最大值选择决定:

$$C_{opt} = \arg \max \{m_{\oplus}(T \in C_i)\} \quad (17)$$

本文所提出方法的流程图如图 4 所示.

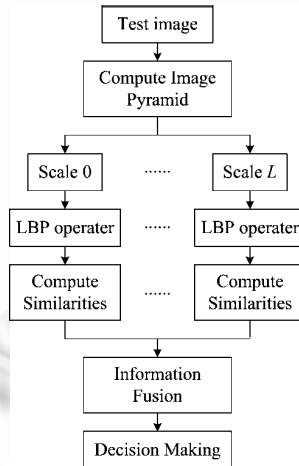


Fig.4 Flowchart of the proposed method

图 4 本文算法的流程图

2 实验结果

为了验证所提出的纹理识别方法,我们在两个常用的纹理图像数据集上完成了相关实验.第 1 个数据集是

Brodatz's album^[19].此数据集包含 112 个不同的纹理图像,每一个图像包含 256×256 个像素.9 种纹理分别取自 Brodatz's album 集合,分别是草、纸、波浪、纤维、沙、木纹、牛皮、羊毛、织物.图 5 显示了初始的 Brodatz's album 图片.第 2 个数据集来自 MIT 视频纹理数据库(VisTex)^[20].我们的实验使用了 40 幅来自此数据库中的纹理图像,图 6 显示了使用的纹理图像.其中的每一幅图像都是拥有 512×512 个像素的自然风景图像.

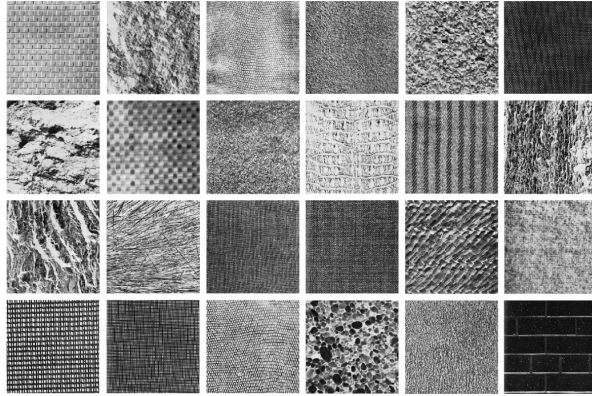


Fig.5 Textures in Brodatz texture database

图 5 Brodatz 纹理数据库中的纹理



Fig.6 Textures in Vistex texture database

图 6 Vistex 纹理数据库

在实验中,仅仅用到通过图像水平分量计算得到的灰度水平.纹理图像通过标准偏差纠正来减少由于灰度变化带来的区别,这是与局部图像纹理无关的部分.每个校正图像的平均灰度值设置为 128,标准偏差为 40.

在本文所提出的方法中,有两个参数是不稳定的.一个是金字塔的层数,另一个是每一种类所使用训练图像的数量.在实验中,对每一种类使用的训练图像保持一个常量.为了验证文中所提出的方法,本文设计了 4 个实验进行验证.第 1 个实验测试在不同图像金字塔层级上对纹理正确识别的效率,第 2 个实验测试对于每一种类拥有不同数量的训练图像的正确识别效率,第 3 个实验测试所提出的方法对于图像旋转的鲁棒性,最后一个实验测试所提出方法的时间消耗.

为了将本文所提出的纹理识别方法与其他方法进行比较,有 3 种以前提出的方法在实验中得到应用:基于灰度共生矩阵(GLCM)的算法^[21]、基于伽柏小波算法^[22]和基于 LBP 算法^[23].GLCM 是一个统计描述影像中的一个局部区域或整个区域相邻像素或一定间距内两像素灰度呈现某种关系的矩阵.该矩阵中的元素值代表灰度级之间联合条件概率密度 $p(i,j|d,\theta)$, p 表示在给定空间距离 d 和方向 θ 时,灰度以 i 为始点,出现灰度级为 j

的概率(也即频数).在本文的实验中,设 $d=1, \theta=0^\circ, 45^\circ, 90^\circ, 135^\circ$. Gabor 滤波器参数设置分别是:方向 θ_0 , 中心频率 (U_l, U_h) , 相位 s 设置如下:

$$\begin{cases} \theta_0 = \left[0, \frac{\pi}{8}, \frac{\pi}{4}, \frac{3\pi}{8}, \frac{\pi}{2}, \frac{5\pi}{8}, \frac{3\pi}{4}, \frac{7\pi}{8} \right] \\ (U_l, U_h) = (0.05, 0.4) \\ s = 6 \end{cases} \quad (18)$$

2.1 对不同金字塔层级效果

使用图像金字塔的目的是扩大初始 LBP 算子的空间区域的支持.理论上,一个更高层数的图像金字塔可以捕获更多的大规模纹理特性.然而,图像金字塔是一种低通滤波器.图像金字塔的最高级别越高,失去的图像细节就越多.事实上,纹理特征在某一相对高层级图像金字塔时将会消失,因此,LBP 算子在这个级别没有影响.此外,层级过高会增加分类的计算成本,而带来的精度提高却几乎为 0.因此,重要的是要评估不同的图像金字塔层级对我们提出的方法的影响.在这个实验中,我们将在这两个数据集的图像分为 16 个非重叠方形块.也就是说,图像的第 1 个数据集,每个图像块有 64×64 像素,而对于第 2 个数据集,每个图像块有 128×128 像素.我们随机选择 10 个图像块为每个图像作为训练样本,其余 6 个图像块作为测试样本.两个数据集的实验结果如图 7、图 8 所示,具体数据见表 1.

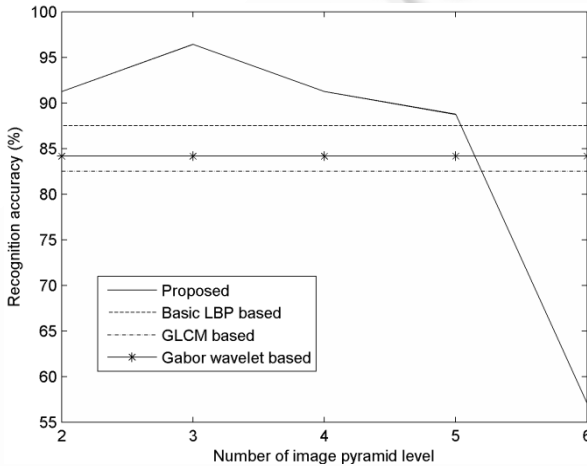


Fig.7 Recognition accuracy of data set 1 by using different image pyramid levels

图 7 对纹理集 1 应用不同金字塔层级的识别精度

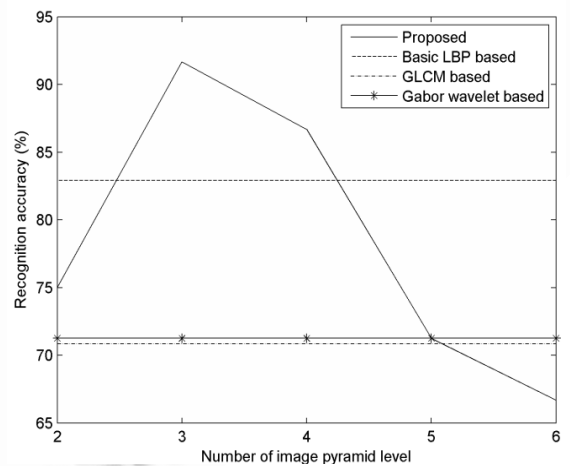


Fig.8 Recognition accuracy of data set 2 by using different image pyramid levels

图 8 对纹理集 2 应用不同金字塔层级的识别精度

Table 1 Recognition accuracy by using different image pyramid levels

表 1 应用不同金字塔层级的识别精度

Method	Levels	Accuracy (%)	
		Data set 1	Data set 2
Proposed	2	91.25	75.00
	3	96.43	91.67
	4	91.25	86.67
	5	88.75	71.25
	6	57.08	66.67
Basic LBP based	—	87.50	82.92
GLCM based	—	82.50	70.84
Gabor wavelet based	—	84.17	71.25

显然,当图像金字塔层数是 2~4 层时,我们的方法优于其他算法.第 1 个数据集,特别是当图像金字塔只有 3

层时,我们的方法达到了最好的识别准确率,为 96.43%,而基本的 LBP 算法、基于 GLCM 的算法以及基于 Gabor 小波算法得到的精度分别只有 87.50%,82.50%和 84.17%。

对于第 2 个数据集,同样,当图像金字塔只有 3 层时,我们的方法达到了最好的准确率,为 91.67%,而基本 LBP 方法、基于 GLCM 的算法以及基于 Gabor 小波算法得到精度分别为 82.92%,70.84%和 71.25%。

值得注意的是,当图像金字塔的层数等于或高于 5 时,识别精度大大降低.出现这样现象的原因是,当图像金字塔的层数太高时,高层图像金字塔所包含的图像细节几乎为 0.因此,LBP 算子不能提取在这些层面的图像金字塔纹理特征,式(12)中会产生错误信念分配.因此,决定信息融合是偏离实际情况。

2.2 不同数量训练样本的影响

训练样本的数量能够对分类器产生很大的影响.训练样本太少不足以捕获类的特性.相反,太多的训练样本将对分类器带来沉重的计算成本.更严重的是,太多的训练样本可能导致过度拟合的问题.在这种情况下,分类器可以准确模拟所有类的不同训练样本之间的关系.然而,分类器是无法推广到未知的新样本,并且测试性能也将急剧恶化.因此,使用不同数量的训练图像测试我们提出方法的性能是值得的。

在这个实验中,所有数据集中的图像被分成 4,16,64 或 256 个非重叠的正方形块,每一块为 64×64.当图像被分成 4,16,64 或 256 块时,开始的第 3,10,50,200 块作为训练样本,其余块分别作为测试样本.在这个实验中,图像金字塔的层数为 3.两个数据集的实验结果如图 9、表 2 所示。

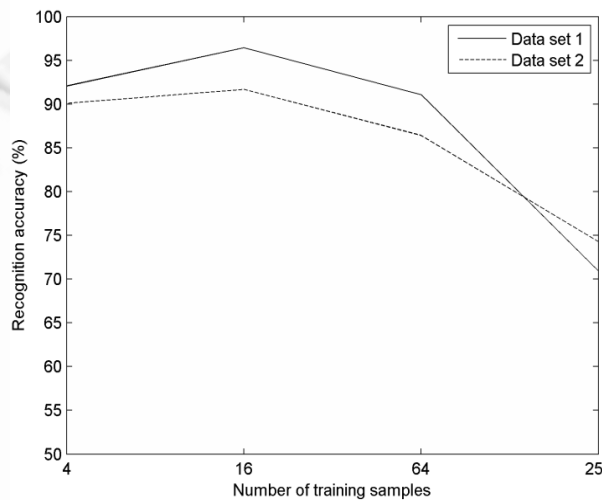


Fig.9 Recognition accuracy by using different number of training samples

图 9 不同数量训练样本的影响

Table 2 Recognition accuracy using different set of training samples

表 2 不同训练样本集 的识别精度值

Data set	Number of training samples	Accuracy (%)
1	4	92.08
	16	96.43
	64	91.07
	256	70.89
2	4	90.10
	16	91.67
	64	86.43
	256	74.29

从两个数据集的数据明显可以看出,将图像分成 4 块或 16 块将产生良好的结果.当图像被分成 64 或 256 块时,识别精度下降.特别是,当图像被分成 256 块时,第 1 个和第 2 个数据集识别精度分别下降到 70.89%和

74.29%.

出现上述现象的最主要的原因是,当只有 3 块作为训练样本时,对当前纹理来说没有足够的信息.如果训练样本的数量太大,两个相邻属于不同类别的样本的距离则通常是很短的,这将导致分类器的性能降低.

2.3 对图像旋转的鲁棒性

真实的纹理可以发生任意旋转.在这个实验中,我们测试了我们的方法对于图像旋转的鲁棒性.实验中,图像金字塔的层数为 3.所有图像数据集被分成 16 个非重叠方形块.第 1 个 10 块作为训练样本.然而,不同于前两次实验,测试样本不能直接来自原始图像.实验中,将原始图像进行一定的旋转,取值为 0~180°之间,步长为 15.两个数据集的实验结果如图 10、表 3 所示.

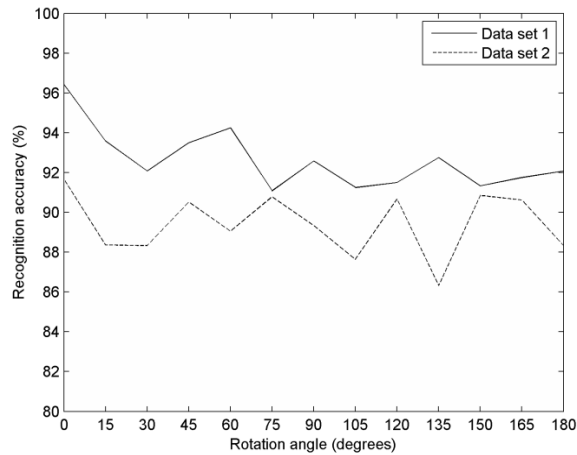


Fig.10 Recognition accuracy of the proposed method under different rotation angles

图 10 不同旋转角度的识别精度

Table 3 Comparison of recognition accuracy under typical rotation angles in this algorithm

表 3 本文算法对典型旋转角度的识别精度的比较

θ	Accuracy (%)	
	Data set 1	Data set 2
0°	96.43	91.67
15°	93.58	88.36
30°	92.08	88.33
45°	93.50	90.52
60°	94.25	89.05
75°	91.08	90.78
90°	92.58	89.33
105°	91.25	87.63
120°	91.50	90.67
135°	92.75	86.33
150°	91.33	90.85
165°	91.75	90.62
180°	92.08	88.33

从结果分析,我们可以得出这样的结论:图像旋转对我们所提出方法的识别精度影响较小.特别是,第 1 个和第 2 个数据集识别精度分别下降不到 5%和 4%.这主要是因为 LBP 算子是一种旋转不变算子^[9].另一方面,高斯滤波器用于图像金字塔也具有旋转不变性.旋转不变性使得我们的方法适合处理真实场景的纹理.

2.4 时间消耗

一些应用需要实时处理,如互动纹理分割和纹理检索.虽然我们的方法不使用数据库来存储纹理特征,但是我们的方法的特征提取步骤仍然是潜在的纹理检索.因此,有必要比较我们的方法与先前提出的其他方法在时

间上的消耗.在这个实验中,每个纹理图像分为 16 块,图像金字塔层数为 3.实验结果见表 4.

Table 4 Time consumption of different methods

表 4 不同方法的时间消耗

Method	Time consumption (s)	
	Data set 1	Data set 2
Basic LBP based	20.48	3.99
GLCM based	140.10	45.11
Gabor wavelet based	104.93	36.50
Proposed	56.81	10.41

显然,基于基本的 LBP 方法需要时间最少,而基于 GLCM 方法需要最多的时间.本文的方法是基于图像金字塔,每个金字塔的上层比更低层的像素少,所以理论上,本文方法的时间消耗几乎是基本 LBP 方法的常数倍,此常数由下式给定:

$$\mu = \frac{1}{4^0} + \frac{1}{4^1} + \dots + \frac{1}{4^{L-1}} = \frac{4^L - 1}{3 \cdot 4^{L-1}} \quad (19)$$

其中, L 是图像金字塔的层数.随着金字塔层数的增加,常量 μ 有一个上限,因为

$$\lim_{L \rightarrow \infty} \mu = \lim_{L \rightarrow \infty} \frac{4^L - 1}{3 \cdot 4^{L-1}} = \frac{4}{3} \quad (20)$$

3 讨 论

本文提出了一种计算简单但有效的基于 LBP 和 D-S 证据理论的多尺度灰度纹理识别方法.基于 LBP 的纹理规则非常简单,但非常有效.这些措施表现很好,尤其是针对 Brodatz 纹理.

根据定义,LBP 算子具有灰度不变性.因此,本文提出的方法是在光照条件变化的情况下有吸引力的,比如在工业检测和遥感应用中.LBP 方法计算简单,相应算子可以实现通过在一个小区域比较几个像素,搜索一个查找表.这使得其实现非常简单、有效.因此,该方法适用于对时间要求较苛刻的应用,比如纹理图像检索.

多分辨率分析可以提高 LBP 方法的性能.为了达到这一目的,本文提出了一种简单的不同空间分辨率的组合方法.实验结果表明,多分辨率分析是有益的,除非单分辨率对于良好的识别力已经足够.

然而,本文的方法有一定的局限性.首先,尽管所提出的方法计算复杂度有一个上界(见第 2.4 节),但该方法的信息融合过程占用了很大一部分时间.这是因为融合几个操作构成,包括乘法和加法.这是为什么本文提出的方法计算时间远远超过基本 LBP 方法.其次,文献[11]中指出,冲突信息的组合将可能导致意外的结果.也就是说,对于本文算法中金字塔的每一个尺度上的信息,如果信息来源是矛盾的,那么判别结果会变得不准确.

对于未来的工作,利用颜色信息的可能性将得到研究.尽管灰度级对大多数纹理分析是有效的,色彩信息通常会带来更多判别准确性.在文献[24]中,一种 LBP 算子可以处理多个颜色通道,然而,还有几个问题需要解决:

- 彩色空间:RGB 颜色空间的共同存储形成颜色图片.但是 RGB 色彩空间不适合分类或分割等图像分析,因为它的通道并不是独立的,而是受光照变化的影响.其他色彩空间,包括 Hue-Saturate-Value(HSV)空间、CIE-Lab 空间、CIE-XYZ 空间等^[25]也不受光照变化影响.但是,哪一种色彩空间最适合基于多尺度 LBP 的彩色纹理识别,并且在不同的应用条件,色彩空间的性能是否稳定还有待调查.
- 颜色通道的处理:与灰度纹理不同,彩色纹理有多个频道.因此,对于式(12),在不同的通道中,信任函数赋值需要重新定义.

另一个潜在的研究兴趣点是处理相互矛盾的信息来源.在我们的方法中,我们使用的是 Dempster 的规范化的组合,但也批评其规范化步骤^[25].因为当信息源发生冲突时,Dempster 的组合规则将导致貌似没有任何意义的结果.文献[26,27]提出了几种组合规则.然而,它们对基于多尺度 LBP 纹理识别是否仍然是行之有效的,还需要进一步去验证.

4 结 论

本文提出了一种基于多尺度 LBP 和 D-S 证据理论的纹理识别方法.它首先计算输入图像的图像金字塔,然后将 LBP 算子应用在图像金字塔的每个尺度上.然后,图像金字塔的每个尺度将产生一个特征向量.通过 D-S 证据理论判别每一尺度上的相似性,多尺度信息通过融合得到组合.

实验结果显示出本文所提出方法的有效性.特别是对于具有 3 层的图像金字塔和为每个纹理提供 16 个训练样本将产生最好的识别精度,在第 1 个和第 2 个数据集上识别精度分别达到 96.43%和 91.67%,分别优于其他算法.同时,本文提出的方法对图像旋转不变性具有一定的鲁棒性.实验结果表明,该方法对两种数据集的识别精度减少不到 5%.此外,本文提出的方法在计算上非常简单,对数据集 1 和数据集 2 的识别分别只需要 56.81s 和 10.41s.

致谢 本研究工作受到国家重点基础研究发展计划(973)(2011CB707904)和国家自然科学基金(61472289)项目组成员的支持和帮助,在此向他们表示衷心的感谢.

References:

- [1] Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans. on Systems, Man and Cybernetics*, 1973,3(6):610–621.
- [2] Conners RW, Harlow CA. A theoretical comparison of texture algorithms. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1980,5(2):204–222.
- [3] Jain AK, Farrokhnia F. Unsupervised texture segmentation using Gabor filters. In: *Proc. of the IEEE Int'l Conf. on Systems, Man and Cybernetics*. Los Angeles: IEEE, 1990. 14–19.
- [4] Ojala T, Pietikäinen M, Harwood D. Performance evaluation of texture measures with classification based on kullback discrimination of distributions. In: *Proc. of the Int'l Conf. on Pattern Recognition*. Jerusalem: IEEE, 1994. 582–585.
- [5] Pietikäinen M, Hadid A, Zhao G, Ahonen T. *Computer Vision Using Local Binary Patterns*. London: Springer-Verlag, 2011.
- [6] Beliakov G, James S, Troiano L. Texture recognition by using GLCM and various aggregation functions. In: *Proc. of the IEEE Int'l Conf. on Fuzzy Systems*. Hong Kong: IEEE, 2008. 1472–1476.
- [7] Vapnik V. *Statistical Learning Theory*. New York: Wiley, 1998.
- [8] Li X, Hu W, Zhang Z, Wang H. Heat kernel based local binary pattern for face representation. *IEEE Signal Processing Letters*, 2010,3(17):308–311.
- [9] Derpanis KG, Wildes RP. Dynamic texture recognition based on distributions of spacetime oriented structure. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. San Francisco: IEEE, 2010. 191–198.
- [10] Derpanis KG, Wildes RP. Spacetime texture representation and recognition based on a spatiotemporal orientation analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2012,34(6):1193–1205.
- [11] Ojala T, Pietikäinen M, Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002,7(24):971–987.
- [12] Dempster AP. Upper and lower probabilities induced by multi-valued mapping. *Annals of Mathematical Statistics*, 1967,3(38): 325–339.
- [13] Shafer G. *A Mathematical Theory of Evidence*. Princetown: Princetown University Press, 1976.
- [14] Kennes R. Computational aspect of the möbius transformation of graphs. *IEEE Trans. on Systems, Man and Cybernetics*, 1992, 2(22):201–223.
- [15] Smets P. The combination of evidence in the transferable belief model. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1990,5(12):447–458.
- [16] Zadeh LA. On the validity of Dempster's rule of combination of evidence. University of California, Berkeley, ERL Memo M79/24, 1979.
- [17] Burt PJ, Adelson EH. The Laplacian pyramid as a compact image code. *IEEE Trans. on Communications*, 1983,COM-31:532–540.

- [18] Wang P, Yang G. Improvement method for the combining rule of dempster-shafer evidence theory based on reliability. *Journal of Systems Engineering and Electronics*, 2005,2(16):471-474.
- [19] Brodatz P. *Textures: A photographic album for artists and designers*. 1966.
- [20] MIT Vision and Modeling Group. *Vision texture*. <http://vismod.www.media.mit.edu>
- [21] Zhou H, Wang R, Wang C. A novel extended local-binary-pattern operator for texture analysis. *Information Sciences*, 2008,7(178): 4314-4325.
- [22] Yuan F. Video-Based smoke detection with histogram sequence of LBP and LBPV pyramids. *Fire Safety Journal*, 2011,3(46): 132-139.
- [23] Beliakov G, James S, Troianov. Texture recognition by using GLCM and various aggregation functions. In: *Proc. of the IEEE Int'l Conf. on Fuzzy Systems (FUZZ-IEEE 2008)*. 2008. 1472-1476.
- [24] Bishop CM. *Pattern Recognition and Machine Learning*. Berlin: Springer-Verlag, 2006.
- [25] Mäenpää T. *The local binary pattern approach to texture analysis—Extensions and applications [Ph.D. Thesis]*. Acta Universitatis Ouluensis C 187, University of Oulu, 2003.
- [26] Tkalcic M, Tasic JF. Colour Spaces: Perceptual, Historical and Applicational Background. In: *Proc. of the IEEE Region 8 EUROCON 2003, Vol.1*. 2003. 304-308.
- [27] Smets P. Resolving misunderstandings about belief functions. *Int'l Journal of Approximate Reasoning*, 1992,3(6):321-344.



孙俊(1972—),男,湖北黄冈人,博士生,主要研究领域为视频目标识别与跟踪。
E-mail: 47325694@qq.com



陈晓(1980—),男,博士生,主要研究领域为图像处理。
E-mail: 84299889@qq.com



何发智(1968—),男,博士,教授,博士生导师,主要研究领域为图形图像处理,协同计算。
E-mail: fzhe@whu.edu.cn



袁志勇(1962—),男,博士,教授,博士生导师,主要研究领域为图形图像处理。
E-mail: yzypcc@163.com