

# 基于FRD的图像纹理情感语义提取

王莉

(山西大同大学数学与计算机科学学院, 大同 037009)

**摘要:** 图像的低层视觉特征(颜色、纹理和形状)中包含大量人类可感知的情感语义信息。利用纹理特征, 提出一种新的索引方法——模糊认识度(FRD)聚类法, 用来描述与情感相关联的语义图像。FRD聚类法能从高层的情感概念出发进行图像检索。索引使用3个感性的纹理特征: 方向性, 对比度和粗糙度生成FRD值。实验采用室内装饰图片, 结果表明, 该方法性能较好。

**关键词:** 模糊认识度; 情感语义; 纹理

## Emotional Semantics Extraction from Image Texture Based on Fuzzy Recognize Degree

WANG Li

(School of Mathematical and Computer Science, Shanxi Datong University, Datong 037009)

**【Abstract】** The low-level visual features(color, texture and shape) in image implicate lots of emotional semantics information with human perception. This paper proposes a new indexing scheme, called Fuzzy Recognize Degree(FRD) clustering using the texture feature. This method allows to retrieve images for emotion-assisted semantic analysis based on high-level affective concepts. For this purpose, a set of perceptual relevant features is introduced: directionality, contrast and coarseness. Collecting indoor decoration images, experimental result shows the high performance of the proposed technique.

**【Key words】** Fuzzy Recognize Degree(FRD); emotional semantics; texture

### 1 概述

纹理是人体视觉系统的一个重要特征, 它揭示了场景的深度和表面的方向, 描述图像区域的性质, 如平滑度, 粗糙度和规律性等<sup>[1]</sup>。纹理特征实际上是颜色特征的另一种表达方式, 有些系统不使用纹理。不过, 在图像分类方面, 纹理提供了重要的信息, 由于它描述了许多反映真实世界图像的内容, 如果皮、云彩、树木、砖头和纤维等。因此在确定高层情感语义时, 纹理发挥着重要作用<sup>[2]</sup>。

为添补情感语义和视觉特征之间的语义鸿沟, 使用适当的先验知识是很重要的<sup>[3]</sup>。本体、用户偏好概况以及相关反馈等技术都用到了这样的知识。但这些方法代价较高, 且只能适用于有限的领域<sup>[4]</sup>。

本文提出一种新的聚类规则——模糊认识度(Fuzzy Recognize Degree, FRD)聚类法, 能保证较高的检索质量。按该规则可对图像进行情感概念的检索, 情感概念由 Tamura 纹理空间得出。情感概念是对人类图像感知的一种语义解释, 被自动抽取表示为 FRD, 存于图像数据库中。例如, 输入情感性查询语句, 如“find cool images”和“find lovely images”, 得到相应情感和纹理信息的图像。

### 2 纹理空间信息

纹理尽管已经受到广泛应用, 但并没有一个被普遍接受的定义。Haindl 指出: 纹理通常是一种表面的视觉属性, 表示了包含在物体表面上的空间信息。相应地, 目前存在许多纹理的定义, 有若干方法来分析纹理特征。

在计算机视力和视觉理解领域, 学者对感性相关特性的纹理描述进行了大量研究。这些研究一般从提出了6个纹理

特征出发, 运用 Tamura 描述的可计算特征度量来验证个人实验的可行性。

本文对 Tamura 纹理的3个感性特性: 方向性, 对比度和粗糙度中加以描述、计算和定义。这3个索引表示在数据库中的“情感标签”, 按一种与人类感知系统一致的方式区分不同的纹理。

#### 2.1 方向性

Haralick 提出对方向性的量度一般基于协方差矩阵。传统分析方法通过坡度值分析分配角度来得出更为优良的方向性(与“各向同性”相对)考量, 这里着重于纹理的一种特定方向支持程度的计算。换句话说, 不去计算“方向”的全局量值, 而是提供一系列“关联支持度”。

假设  $T$  表示  $n \times m \times c$  颜色个数的纹理;  $v=(x, y)$  表示偏移向量; 协方差矩  $C(T, v)$  是一个  $c \times c$  维矩阵, 表示如下:

$$C(T, v)_{ij} = |\{(p, q) \text{ in } T \times T : q = p + v \text{ and } L(p) = i \text{ and } L(q) = j\}| \quad (1)$$

其中,  $L(p)$  是像素  $p$  的亮度值;  $T$  中分别包含灰度值  $i$  和  $j$  以及置换向量  $v$ ;  $C$  中点  $(i, j)$  包含  $T$  中的像素对的数量。

根据 Julesz 推测, 协方差矩阵收集纹理的二阶统计数据, 用来区分不同纹理之间的特征。在分类和检索技术中, 这一方法已经受到了广泛应用。基本原理是, 当偏移向量  $v$  与相应方向相关时, 包含  $v$  的协方差矩阵的主对角线结构更接近

**基金项目:** 山西大同大学青年研究基金资助项目(2008Q14)

**作者简介:** 王莉(1982-), 女, 助教、硕士, 主研方向: 智能信息处理

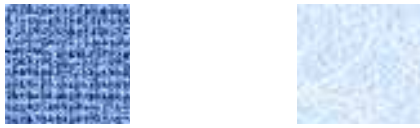
**收稿日期:** 2009-02-20      **E-mail:** joininwangli@163.com

于图像的直方图。从这一原理出发，首先对特定方向相关度进行估计，再计算出有多少对角线协方差矩阵与偏离出图像直方图的该方向相关联。具体方法如下：

假设协方差矩阵  $C(T, v)$ ,  $d_c(t) = C(T, v)_{tt}$ ,  $t = 0, 1, \dots, c$ ,  $H(t)$ 为  $T$ 的直方图。计算直方图和  $d_c$ 之间的结构区域，结果作为协方差矩阵和直方图之间的偏离度。矩阵  $C$ 的偏差定义如下：

$$d(c) = \sum_{t=2}^{c-1} [H(t) - d_c(t)] + \frac{1}{2} \times [H(1) - d_c(1) + H(c) - d_c(c)] \quad (2)$$

为从有限偏移集  $S$  中获得上式中定义的偏差测量的方向相关度，采用以下方法： $D_S$ 是与  $S$  偏移有关的协方差矩阵得出的偏差值矩阵， $d_s$ 是  $D_S$ 在  $[0, 1]$ 范围内规格化的数值 (levels  $d=1, 2, \dots, D$ )集合。集合  $R_s = 1 - d_s$  包含  $S$  中方向的相关度。图 1 显示了 2 种纹理结构，它们具有不同方向性特征。



(a)Fabric 8, *directionality*=0.897 45 (b)Rock 5, *directionality*=0.439 25

图 1 方向性数值不同的纹理

## 2.2 对比度

局部对比度通常定义为，每个像素点对周围像素点变化的估计。假设像素点  $p=(i, j)$ 以及周围点分布  $W \times W$ ，局部对比度定义如下：

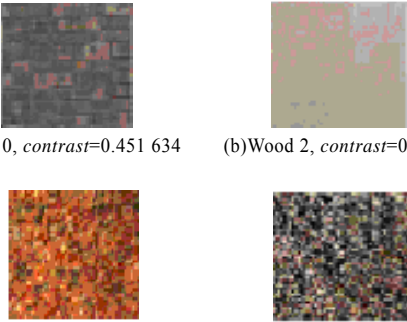
$$local\_contrast(i, j) = \frac{\max_{p \in W \times W} (p) - \min_{p \in W \times W} (p)}{\max_{p \in W \times W} (p) + \min_{p \in W \times W} (p)} \quad (3)$$

全局对比度是局部对比度的算术平均值，计算如下：

$$contrast = \frac{1}{m \times n} \times \sum_{i=1}^n \sum_{j=1}^m local\_contrast(i, j) \quad (4)$$

其中， $n, m$ 分别是图像的长和宽。

图 2 显示了 VisTex 数据库中的 4 种纹理结构，它们具有不同的对比度特征。



(c)Fabric 15, *contrast*=0.652 579 (d)Level2 12, *contrast*=0.797 274

图 2 对比度数值不同的纹理

## 2.3 粗糙度

粗糙是最重要的感性纹理特性，许多情况下词汇“粗糙”可以代替“纹理”来使用。如果 2 个纹理仅仅在大小上有不同，那么尺寸大的更粗糙。

根据 WordNet 定义粗糙度：“与较大微粒合成度的质量描述”。Rosenfeld 提出一种粗糙度的计算公式，使用不同大小的算子对结构化纹理元素大小的计算，具体步骤如下：

(1)收集  $K$  幅图像，每幅图像每个像素的亮度都是周围像素  $2^k \times 2^k$  ( $k=1, 2, \dots, K$ )的亮度均值：

$$A_k(x, y) = \frac{1}{2^{2k}} \times \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} T(i, j) \quad (5)$$

(2)分别计算非相邻点对在垂直和水平方向上的均值差：

$$E_{k\_horizontal}(x, y) = |A_k(x+2^{k-1}, y) - A_k(x-2^{k-1}, y)| \quad (6)$$

$$E_{k\_vertical}(x, y) = |A_k(x, y+2^{k-1}) - A_k(x, y-2^{k-1})| \quad (7)$$

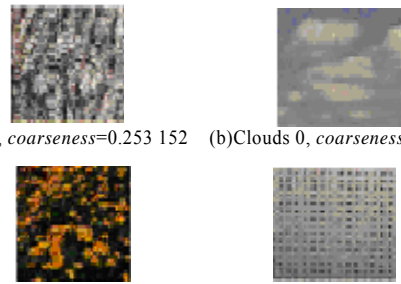
(3)计算每个像素点上周围像素点(结构元素)的最佳卷积大小，作为相邻区域的最大不同点。 $S_{best}(x, y) = 2^k$ ,  $k$ 使不同达到最大：

$$E_k = \max(E_{k\_horizontal}, E_{k\_vertical})$$

(4)计算  $S_{best}$  的算术平均值，作为全局粗糙度：

$$coarseness = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n S_{best}(i, j) \quad (8)$$

为获得粗糙度的测量值，须将  $S_{best}$  的值在  $[0, 1]$ 内标准化。图 3 显示了 4 种纹理结构，具有不同的粗糙度特征。



(c)Flower 0, *coarseness*=0.457 542 (d)Tile 7, *coarseness*=0.753 142

图 3 粗糙度数值不同的纹理

## 3 纹理的 FRD-群度

FRD-群的框架和执行过程见图 4。

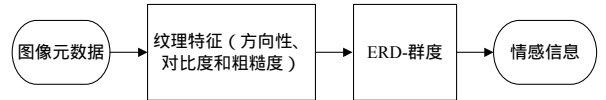


图 4 FRD-群的框架和执行过程

该执行过程先从图像中提取出 Tamura 纹理特征并存放入数据库，这些已存的 Tamura 纹理特征被转换成 FRD 群度。FRD 群度被存放在一张情感类别表中。FRD 群度用来从情感概念角度搜索图像。

### 3.1 纹理的高层语义

FRD 群表示情感词汇和纹理近似语义之间的概念性距离，前者由模糊认知函数生成，后者的原理基于 Tamura(定向性、对比度和粗糙度)纹理空间。表 1 显示 10 类由 Tamura 纹理特点得出的情感词汇。

表 1 纹理中的近似语义

类	语义
1	homogeneous/non homogeneous
2	geometrical/non geometrical
3	soft/rough
4	fine/coarse
5	flat/non flat
6	regular/irregular
7	symmetrical/non symmetrical
8	simple/complex
9	defined/diffuse
10	natural/artificial

根据方向性不同, 图像可能给人一种规则或不规则的感觉。根据对比度不同, 图像可能给人一种柔和或生硬的感觉。另外, 根据粗糙度的差异, 图像可能给人一种丰富或单调的感觉。人类观察图像之后得出的感觉往往差异巨大, 而且有时是模棱两可的。某些纹理可以很好地区分它们的差异, 用来进行分类效果明显。例如方向性, “规则的”或“平滑的”感觉的可以有效地区分。因此, 情感词汇和纹理之间的 FRD 群度可由下式得到:

$$Frd(f, a) = \max\{\min[Frd(f, t_i), Frd(t_i, a)], 0\} \quad i = 1 \dots 11 \quad (9)$$

由纹理特征得出的情感词汇空间, 表示词汇的认知程度, 被作为模糊测量度存入数据库中。表 2 显示了纹理和情感词汇之间的相关联测量度, 用来计算 FRD 群度。

表 2 纹理和情感词汇的关联测量度

Affective terms	柔和度	清晰度	规则性
活跃的	0.98	0.90	0.50
强烈的	0.90	0.70	0.48
浪漫的	0.70	0.90	0.8
有节奏的	0.80	0.88	0.65
美丽的	0.75	0.90	0.65
动人的	0.78	0.70	0.60
帅气的	0.64	0.75	0.65
可爱的	0.70	0.90	0.80
活泼的	0.65	0.85	0.78

如果  $f$  与纹理的情感词汇有关, 那么包括在  $f$  中的情感  $a$  的可能性可以通过计算得出。例如, 假设关于“柔和”和“清晰”的情感词汇模糊集是: 柔和= $\{(0.98, \text{活跃的}), (0.90, \text{强烈的}), (0.70, \text{浪漫的}), (0.75, \text{美丽的}), (0.70, \text{可爱的})\}$ ; 清晰= $\{(0.90, \text{活跃的}), (0.70, \text{强烈的}), (0.90, \text{浪漫的}), (0.90, \text{美丽的}), (0.90, \text{可爱的})\}$ ; 而“有生气的”情感词汇模糊集是: 有生气的= $\{(0.98, \text{柔和}), (0.95, \text{清晰}), (0.80, \text{规律性}), (0.50, \text{精良}), (0.10, \text{平滑}), (0.90, \text{几何形})\}$ 。由式(9)可知, “活跃的”和“有生气的”的 FRD 群度计算如下:

$$Frd(\text{dynamic}, \text{active}) = \max\{\min[Frd(\text{dynamic}, \text{softness}), Frd(\text{softness}, \text{active})], \min[Frd(\text{dynamic}, \text{clearness}), Frd(\text{clearness}, \text{active})]\} = \max\{\min(0.95, 0.98), \min(0.90, 0.98)\} = 0.95$$

可以得出结论, 纹理特征“柔和度”或“清晰度”较高的图像给人以“有生气的”感觉, 而“活跃的”情感包含在“有生气的”情感中的关联可能性达到了 0.95。所以可知, “柔和度”或“清晰度”较高的图像带有“活跃的”情感。

### 3.2 FRD-群度的生成

根据方向性的不同分布, 将图像纹理的语义进行分组。FRD 群度存入数据库中, 用来完成分类工作。依据方向性, Tamura 纹理空间分布在平面的  $360^\circ$  范围内。这里将纹理分为 10 个类别, 来代表由人类观察中的纹理。假设“柔和的”的方向性程度范围为:  $15 \leq d \leq 360 \leq 345$ 。10 种纹理根据对比度不同将会变得更加昏暗或是明快, 根据粗糙度不同将会变得更加柔和或是生硬。例如, 如果图像  $I_n$  的方向性、对比度、粗糙度程度 ( $direction, contrast, coarseness$ ) 是 (5, 250, 30), 方向性程度处于  $15 \leq d \leq 360 \leq 345$  范围内, 便可得出  $I_n$  是“柔和的”图像的结论。值得注意的是, 根据粗糙度的不同, “柔和的”特征可能会减少甚至消失。如果粗糙度处于  $20 \leq con \leq 0$  范围内, 即使方向性程度范围为:

$15 \leq d \leq 360 \leq 345$ , 也不可能获得“柔和的”纹理。

由此可判断, 在人们看到纹理感受情感时, 对比度和粗糙度扮演了关键的角色。另外, 当方向性的数值发生变化时, 对图像纹理的感觉会变得模糊。例如, 如果方向性的数值为 0, 纹理的区分是清楚的。如果方向性的数值处在 15~20 之间, 想要区分纹理可能会变得十分困难。鉴于此, 当纹理发生变化时, 一个范围在  $\pm 5$  的容错值将会引入。图 5 为利用 Tamura(方向性、对比度和粗糙度)空间自动生成 FRD 群度, 并存入数据库的算法流程。

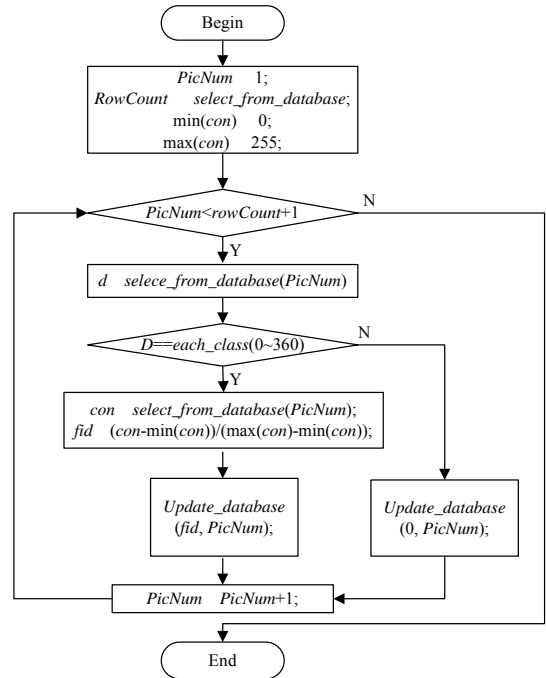


图 5 FRD 群度的生成算法流程

### 3.3 基于 FRD 群度的图像索引和检索

每个图像都具有纹理特征和情感特征, 对图像的处理也包括这两个方面。FRD 群度由 FRD 群算法直接生成, 根据预定义的 12 种情感分组特性, 决定图像应该属于哪个类别。图 6 显示了将图像数据存入情感分组的过程。

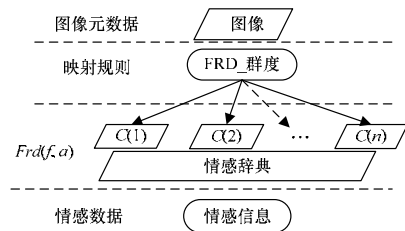


图 6 通过情感概念对图像进行分类和检索的过程

每个情感分组建立一个能够支持情感形容词的情感辞典, 来搜索与术语辞典和感动形容词有关的图像。假设输入“find active images”, 首先扫描情感形容词。如果没有适当的形容词, 一个近似的形容词将被启用, 作为情感形容词。这一情感形容词用来计算模糊认知度, 得出 FRD 群度, 然后在一张情感分组表中查找相同的 FRD 群度, 最终得出图像的检索结果。

## 4 实验结果及分析

在实验中, 设计若干数据库来评估 FRD 群的检索性能, 研究 FRD 本身特征, 并将 FRD 的检索性能与其他传统方法进行比较。实验数据来自麻省理工学院视觉纹理数据库

(VisTex)。任意选择 200 个室内装饰图片作为检索对象,描述不同纹理的特征。

情感概念图像检索技术,采用基于相似性的匹配技术来代替确切性的匹配技术。实验利用查全率和查准率来评价实验结果的性能。查全率指被检索出的相关结果数量在检索系统相关数据总量中所占的比例。查准率指被检索出的相关结果数量在被检索数据总量中所占的比例。

实验对 8 幅待查图片在 VisTex 数据库中分别利用 FRD 技术和 Tamura 技术进行查找,并根据查全率和查找时间对 2 种方法进行比较。结果如表 3 所示,可见,FRD 技术在查全率上均匀高于 Tamura 技术,并在查找时间上又均少于 Tamura 技术。

表 3 Tamura 和 FRD 的性能对比

图片编号	查全率		查找时间/s	
	FRD 技术	Tamura 技术	FRD 技术	Tamura 技术
Pic1	0.602	0.558	0.680	0.790
Pic2	0.809	0.697	0.700	0.810
Pic3	0.887	0.456	0.380	0.710
Pic4	0.768	0.503	0.310	0.690
Pic5	0.492	0.413	0.220	0.410
Pic6	0.708	0.653	0.230	0.430
Pic7	0.715	0.698	0.250	0.700
Pic8	0.733	0.728	0.190	0.450

## 5 结束语

本文介绍基于图像纹理的 FRD 群概念、生成算法以相关分组技术,可以通过输入情感概念词汇检索图像。FRD 群由可感知的 Tamura 纹理空间,如方向性、对比度和粗糙度生成。情感形容词依附于图像和模糊度,模糊度通过人类视觉特征

自动获得。感动形容词向图像和模糊的程度被运用自动地被获得根据人的视觉特点。FRD 群检索规则支持源自人类感觉和情感的基于语义的图像检索,也支持基于视觉特征的传统检索技术。根据情感表达,情感概念被分为 12 类,所有图像也被分入 12 类中。为每个类别设置合适的 FRD 群度,图像的检索速度可以得到提高。可以预见,以情感表达为核心的更加用户友好和更加准确的检索技术一定能够实现。

进一步研究方向是合适地对纹理信息进行分类,寻找更加有效的分类方法,以便提高 FRD 群的检索效率。同时,为了保证更高检索质量,研究关于怎样采用图像的其他视觉特点,在 FRD 群的基础上运用多特征组合譬如颜色和形状,对基于情感的图像检索具有广泛的研究价值和意义。

## 参考文献

- [1] Battiato S, Gallo G, Nicotra S. Perceptive Visual Texture Classification and Retrieval[C]//Proceedings of the 12th International Conference on Image Analysis and Processing. Mantova, Italy: [s. n.], 2003: 17-19.
- [2] Liu Ying, Zhang Dengsheng, Lu Guojun, et al. A Survey of Content-based Image Retrieval with High-level Semantics[J]. Pattern Recognition, 2007, 40(1): 262-282.
- [3] 王伟凝, 余英林. 图像的情感语义研究进展[J]. 电路与系统学报, 2003, 8(5): 101-109.
- [4] 王上飞, 王煦法. 基于情感模型的感性图像检索[J]. 电路与系统学报, 2003, 8(6): 48-53.

编辑 索书志

(上接第 209 页)

- [5] 刘晨华, 冯象初, 张力娜. 基于离散小波阈值的偏微分图像去噪[J]. 计算机工程, 2008, 34(15): 196-198.
- [6] Malfait M, Roose D. Wavelet-based Image Denoising Using a Markov Random Field a Priorimodel[J]. IEEE Trans. on Image Processing, 1997, 6(4): 549-565.
- [7] Shi Rongjie, Shen Ifan, Chen Wenbin B. Image Denoising Through

Locally Linear Embedding[C]//Proc. of Computer Graphics, Imaging and Vision. New Trends, Beijing, China: [s. n.], 2005: 147-152.

- [8] Decoste D. Visualizing Mercer Kernel Feature Spaces via Kernelized Locally-linear Embeddings[C]//Proc. of ICONIP'01, Shanghai, China: [s. n.], 2001.

编辑 金胡考

(上接第 211 页)

## 4 结束语

本文通过在梯度域上对图像动态范围进行压缩调整,使得高动态范围的图像能够在常规显示设备上得到良好的显示,使图像在观察时和真实场景一致。从对 HDR 图像进行处理的过程和结果来看,该方案适用于诸如数字摄影、电影艺术、科学图像增强等要求高质量静态图像输出的情况。

## 参考文献

- [1] Larson G W. A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes[J]. IEEE Transactions on Visualization and Computer Graphics, 1997, 3(4): 291-306.
- [2] Tumblin J, Lics T G. A Boundary Hierarchy for Detail-preserving

Contrast Reduction[C]//Proc. of SIGGRAPH'99. Los Angeles, California, USA: [s. n.], 1999.

- [3] Durand F, Dorsey J. Fast Bilateral Filtering for the Display of High Dynamic Range Image[C]//Proc. of SIGGRAPH'02. [S. 1.]: ACM Press, 2002: 257-265.
- [4] Fattal R, Lischinski D, Werman M. Gradient Domain High Dynamic Range Compression[J]. ACM Transactions on Graphics, 2002, 21(3): 362-371.
- [5] Demmel J, Berkeley U C, Page H. Applications of Parallel Computers[EB/OL]. (1996-03-21). <http://www.cs.berkeley.edu/~demmel/cs267/lecture24/lecture24.html>.

编辑 陆燕菲