

一种基于感兴趣内容的图像检索方法

沈项军, 崔志明, 龚声蓉, 马小虎

(苏州大学计算机科学与技术学院, 苏州 215006)

摘要: 在基于内容图像检索研究中, 针对如何提取用户感兴趣的内容以提高检索效果的问题, 该文根据用户选择的相似图像自动获取“感兴趣”区域, 使用 Fisher 判别进行区域中“感兴趣”特征分量阈值的动态选择, 并以该“感兴趣”特征分量及“感兴趣”区域作为用户感兴趣的内容进行图像检索。试验结果表明了该方法的有效性。

关键词: 动态阈值选择; “感兴趣”特征分量; 基于内容图像检索

Image Retrieval Method Based on Interesting Contents

SHEN Xiang-jun, CUI Zhi-ming, GONG Sheng-rong, MA Xiao-hu

(School of Computer Science and Technology, Soochow University, Suzhou 215006)

【Abstract】 In the research area of Content Based Image Retrieval(CBIR), aiming at extracting the interesting image content to improve retrieval results, this paper proposes an algorithm based on dynamical region pair matching through selecting images by users, uses the Fisher prediction to obtain the dynamical threshold to explore the “interesting” components of the features in the region pair match, adopts those interesting components of the features and the region pairs to retrieve the similar images. The experimental results show the method is effective.

【Key words】 dynamical threshold selection; “interesting” components of features; content-based image retrieval

如何利用特征的相似性来提高检索的效果, 是基于内容的图像检索领域关心的一个问题, 在此问题上, Li等^[1]提出了一种“感兴趣”特征分量选择方法, 并提出一种不同于 Minkowski 距离的距离度量方法(DPF)。该距离计算方法通过阈值判断, 将特征向量中的各个分量聚类成“感兴趣”和“不感兴趣”特征分量, 在图像检索中只需计算“感兴趣”特征分量的距离。然而 Li 等并没有提出一种自动选择阈值的方法, 而是采用的固定阈值。不同的图像语义应该有不同数量的“感兴趣”特征分量, 这不符合人的心理认知。在图像层面上, 作为检索范例的图像, 感兴趣的可能只是某些区域或是图像中的某些对象。在“感兴趣”区域的提取上, 通常采用人机交互的方法, 这显然增加了用户的负担, 需要找到一种能自动学习用户“感兴趣”区域的方法。本文提出了一种动态选择“感兴趣”特征和“感兴趣”区域匹配的方法: (1) 根据用户选择的图像, 利用图像相似的空间结构特性, 学习图像间的相似区域, 并剔除掉匹配错误的区域, 得到的这些相似区域对被视为“感兴趣”的区域对; (2) 使用 Fisher 准则得到的动态阈值来区分区域匹配对中的“感兴趣”特征分量, 并调整“感兴趣”区域中各个特征的权重和区域的权重, 动态地检索感兴趣图像内容。

1 区域匹配对算法

为了获得图像的区域信息, 需要对图像进行分割处理, 笔者采用了 Deng 等人提出的图像分割算法^[2], 并对分割的图像区域进行配对。

1.1 图像特征

在进行区域配对之前, 需要介绍使用特征, 笔者选择了 3 种特征进行区域配对和“感兴趣”特征分量的选择: 色彩, 纹理, 形状特征。为了加快检索的速度, 采用了二级特征选择策略——统计粗尺度特征和直方图细尺度特征, 各类特征

如下: 在色彩特征的选择上, 粗尺度特征选择区域的色彩均值向量; 而在细尺度特征上, 选择文献[3]提出的动态色彩直方图和直方图距离计算方法。

在纹理特征的选择上, 粗尺度特征上选择 LBP 直方图^[4]。它的取值范围为 0~255。区域 LBP 直方图 HT 表达为

$$HT(m) = n_m/N, m=0,1,\dots,255 \quad (1)$$

其中, n_m 是区域中 LBP 值为 m 的像素个数; N 是区域中的像素总数。

对纹理粗尺度特征采用计算 LBP 直方图的统计特征:

$$(1) \text{能量} := \sum_{i=0}^{255} HT(i)^2 \quad (2)$$

$$(2) \text{方差} := \sum_{i=0}^{255} (HT(i) - 1/256)^2 \quad (3)$$

$$(3) \text{熵} := -\sum_{i=0}^{255} HT(i) \log(HT(i)) \quad (4)$$

在最后一个形状特征的选择上, 粗尺度的形状特征选择文献[5]介绍的重心及其形状二阶矩; 在细尺度特征上选择文献[6]中的直方图。

纹理和形状直方图都采用下式计算其相似距离, 即

$$DH_i = \sum_j (HIST_{i,1}(j) - HIST_{i,2}(j))^2 \quad (5)$$

其中, j 代表维数; $H_{i,1}, H_{i,2}$ 代表 2 幅图像直方图特征; $i=s$ 代表计算纹理直方图; $i=t$ 代表计算形状直方图。

粗尺度特征或是细尺度特征距离的计算公式为

$$Dist = W_c DC(x,y) + W_t DH_t + W_s DH_s \quad (6)$$

其中, W_c 是色彩特征的权重; W_t 是纹理特征权重; W_s 是形状特征权重; $DC(x,y)$ 是文献[3]引入的色彩向量 x,y 的动态直方

基金项目: 国家自然科学基金资助项目(60673092)

作者简介: 沈项军(1977 -), 男, 博士, 主研方向: 模式识别, 图像检索; 崔志明, 教授、博士生导师; 龚声蓉、马小虎, 教授

收稿日期: 2007-02-22 **E-mail:** xjshen@suda.edu.cn

图距离； DH_t 和 DH_s 定义同上。同时，未进行区域匹配学习前各个特征的权重按照经验获得： $W_c=0.45$, $W_l=0.3$, $W_s=0.25$ 。

1.2 区域匹配对算法

1.2.1 获得初始区域匹配对

假设用户选择 N 幅图像进行学习，为减少计算开销，并考虑到用户选择的图像之间在语义上具有相关性，在上述图像中选择分割区域最多的图像作为参考图像；剩余的 $N-1$ 幅图像作为待匹配图像。设参考图像为 I_a 、待匹配图像为 I_b ， I_a 中的第 i 个区域、 I_b 中的第 j 个区域以及这两个区域之间的相似距离分别用 R_a^i 、 R_b^j 和 $SD(R_a^i, R_b^j)$ 标记。 $SD(R_a^i, R_b^j)$ 的值根据相应区域的色彩、纹理和形状特征，采用式(6)计算得到。如果 R_a^i 、 R_b^j 之间同时满足如下3个条件，则将它们视作一个初始区域匹配对：

$$\begin{aligned} SD(R_a^i, R_b^j) &= \min_{x_m} K SD(X_m, R_b^j) \quad X_m \in I_a \\ SD(R_a^i, R_b^j) &= \min_{y_n} K SD(R_a^i, Y_n) \quad Y_n \in I_b \\ SD(R_a^i, R_b^j) & \leq Threshold \end{aligned} \quad (7)$$

其中， $Threshold$ 是系统设定的一个阈值； Min_K 是极小化算符，它将相关的相似距离中满足给定条件的最小的 K 个区域对作为候选的区域配对。在所进行的实验中， $K, Threshold$ 取2, 0.47。若2个区域 R_a^i, R_b^j 相互匹配，记作 $R_a^i \leftrightarrow R_b^j$ 。

1.2.2 删除错误区域匹配对

考虑由参考图像 I_a 和待匹配图像 I_b 中的全部初始区域配对所构成的集合 MP ，即

$$MP = \{ mp_i | R_a^i \leftrightarrow R_b^d, d=1, 2, \dots, k \} \quad (8)$$

其中， R_a^i 是 I_a 中的第 i 个区域； R_b^d 是 I_b 中与 R_a^i 匹配的第 d 个区域。

设参考图像中的一个区域在相应的待匹配图像中有 k (K)个与之具有配对关系的区域。显然，这 k 个初始区域配对并不一定都是正确的区域配对，其中一些可能是错误的。

本文使用图像的空间结构特性来删除这些错误的区域配对。设 $R_a^i \leftrightarrow R_b^m$ 和 $R_a^j \leftrightarrow R_b^n$ 分别是2个满足配对条件的初始区域配对，所涉及的区域 $R_a^i, R_a^j, R_b^m, R_b^n$ 的重心分别用

$$(G_{a,x}^i, G_{a,y}^i), (G_{b,x}^m, G_{b,y}^m), (G_{a,x}^j, G_{a,y}^j), (G_{b,x}^n, G_{b,y}^n)$$

当 R_a^i 和 R_a^j 的重心所构成的向量 $A = (G_{a,x}^i - G_{a,x}^j, G_{a,y}^i - G_{a,y}^j)$ 以及由 R_b^m 和 R_b^n 的重心所构成的向量 $B = (G_{b,x}^m - G_{b,x}^n, G_{b,y}^m - G_{b,y}^n)$ 之夹角的值小于指定的阈值时，认为相应的两个初始区域配对是正确的区域配对。这里，隐含地要求正确的区域配对在空间结构上应满足所谓的同构性，即要求各图像所含区域之间的相对位置关系应大致相同。

在此基础上，进一步考虑含有多个区域的图像之间的区域配对问题。显然，正确的区域配对在所有的情况下都应该保持良好的同构性，反之则不然。为此，设置了两个计数器 MC 和 UMC 。其中， MC 用于记录 MP 配对集合中与某一区域配对相关联的同构匹配计数， UMC 则用于记录相应的异构错误计数。具体的算法如下：

(1)在 MP 集合中选择相似距离最大的区域匹配对作为最大可能错误区域匹配对。假设区域对为 $\{ r_a^i \leftrightarrow r_b^s \}$ ，其他在 MP 集合中的区域匹配对都是正确的匹配对。

(2)计算与其他 MP 集合中的区域匹配对的重心向量夹角差，如果两个区域匹配对的重心的夹角差大于给定阈值，例如取 $\pi/2$ ，就在 UMC 计数器累加上一定的权重，否则就在

MC 上累加权重。考虑到相邻的区域对在匹配中的重要性，根据同一图像中两个区域是否相邻，分别赋予不同的权：区域相邻，权重为1.2；区域不相邻，权重为1。

(3)计算计数器 MC 在 MC 和 UMC 总和中的百分比。如果百分比大于给定阈值，例如取0.6，说明 $\{ r_a^i \leftrightarrow r_b^s \}$ 与其他区域对能匹配的个数较多，则认为这个匹配对是一个“好”的区域匹配对，需要保留在 MP 集合中；否则，应将该区域匹配对剔除出 MP 集合中。回到步骤(1)中，继续寻找错误区域匹配对，直到 MP 集合中所有的区域匹配对都是符合条件的匹配对为止。

2 “感兴趣”特征动态选择算法

Li等提出的DPF“感兴趣”特征选择函数^[1]，定义为

$$D(m, r) = \left(\sum_{\delta_i \in \Delta_m} \delta_i^r \right)^{\frac{1}{r}} \quad (9)$$

其中， $D(m, r)$ 是以 m, r 为参数的特征距离表达； $\delta_i = |x_i - y_i|$, $i = 1, 2, \dots, t$ 。 $\Delta_m = \{\delta_1, \delta_2, \dots, \delta_m\}$ ， δ_i 是第 i 个分量的特征距离； x_i 和 y_i 是两幅图像的特征向量 X, Y 的第 i 个分量； Δ_m 是前 m 个特征距离最小的特征分量的集合； r 是Minkowski距离测度， $r=2$ 就是欧式空间距离； m 值就是固定阈值。

文献[1]采用的阈值 m 是根据实验获得，不符合实际要求，文中根据得到的区域匹配对，动态的获得其中“感兴趣”特征分量。为此，采用了Fisher模式判别^[7]方法，并对传统的Fisher判别作了改进，以获得最佳分类效果，从而得到动态阈值：

$$B(i) = \frac{|f_1(i) - f_2(i)|}{v_1^2(i) + v_2^2(i)}, \quad i = 1, 2, \dots, t \quad (10)$$

其中：

$$f_1(i) = \frac{1}{i} \sum_{j=1}^i \delta_j \quad (11)$$

$$f_2(i) = \frac{1}{t-i} \sum_{j=i+1}^t \delta_j \quad (12)$$

$$v_1^2(i) = \sum_{j=1}^i |\delta_j - f_1(i)|^2 \quad (13)$$

$$v_2^2(i) = \sum_{j=i+1}^t |\delta_j - f_2(i)|^2 \quad (14)$$

因此，阈值 m 就是在 $B(i)$ 最大处的 i 值，即

$$m = \operatorname{argmax} B(i) \quad (15)$$

由于参考图像的区域在多个待匹配图像中有多个匹配区域，因此 δ_i 是特征第 i 个分量的距离均值。并且在使用式(10)之前，按照均值距离的大小将 δ_i 进行了排序。 argmax 是取最大值的函数， m 取 $B(i)$ 值最大处的 i 的值。

如果特征向量中保留的分量个数占特征维数的比率越大，说明这个特征就越是“感兴趣”的特征，越需要予以保留，则需要动态调整区域中各个特征整体的权重。采用的策略是赢家通吃的规则：即在色彩，纹理和形状特征中，将调整出的权重全都分配给最“感兴趣”的特征。

设 w_1 和 w_2 是不太“感兴趣”特征初始权重， w_3 是最“感兴趣”特征权重。 t_1, t_2, t_3 是相应的特征选择后保留特征分量数在特征维数空间中的比率。权重调整算法如下：

```
For k: = 1 to 2
LW: = (1-t_k)w_k;
w_k := t_k * w_k;
w_3 := w_3 + LW;
End
```

LW 是根据特征维数下降的个数与特征维数的比率乘上该特征原有的权重而得到的该特征损失的权重。

3 实验结果

在获得区域特征的权重调整以后，以匹配区域的面积百分比来得到各个区域匹配对的权重：区域的面积越大，表示区域越重要。衡量图像相似性采用

$$ID = \sum_k w_k d_k \quad (16)$$

其中， ID 是图像的相似距离； w_k 是第 k 个区域匹配对的权重； d_k 是“感兴趣”特征计算得到的区域的相似距离。系统以单个图像进行检索时，也是采用式(16)，不过 w_k 代表图像的第 k 个区域的权重，这样计算的是整幅图像所有区域的相似距离总和。

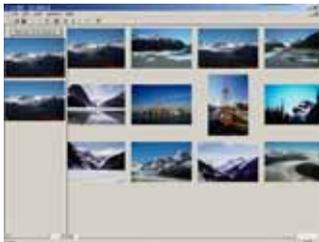
由于是多区域的匹配检索，因此计算量很大。为了降低计算量，加快检索的速度，采用了2级检索策略：计算区域的统计特征，并设定阈值为0.3；计算出的统计特征距离如果大于阈值，就直接舍弃该区域，否则才会进一步比较区域的直方图特征，阈值设定为0.47。保留直方图距离小于阈值的区域，否则就再一次舍弃该区域。

在 Win2000 系统，CPU 为 AMD 2500+，512 MB 内存的计算机中，使用 VC6 和 SQL 2000 开发的 LUSS 系统，检索由 1 400 幅 COREL 图像组成的数据库。这些图像由 14 个大类的图像组成，每类中各由 100 幅预先标注相似的图像组成。

根据上述方法构建的系统 LUSS。不同区域学习检索结果见图 1。落日语义的检索结果见图 2。



(a) 未经学习的 1 幅图像的检索结果



(b) 经过区域学习的 2 幅图像的检索效果

图 1 不同区域学习检索结果

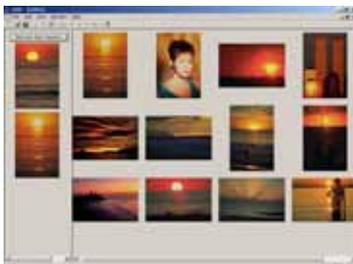
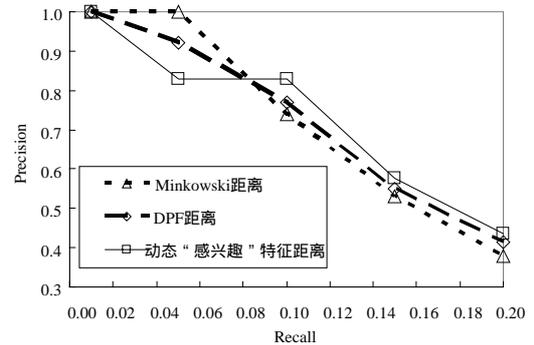


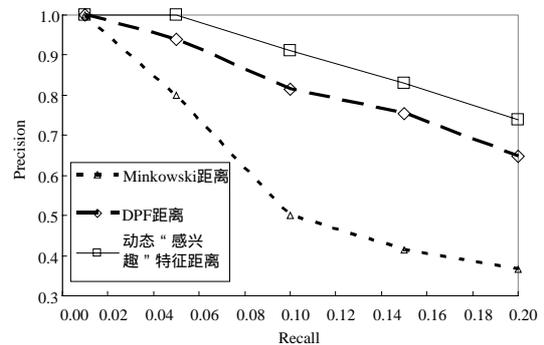
图 2 落日的区域检索

从图 1 中可以看到，用户选择了一幅没有经过学习的图像的检索，前 12 个最相似的结果中，有 5 个是错误的检索结果；经过 2 幅图像自学习后的检索，错误个数降低到 2 个。从而得到更好的语义检索效果。从图中看到，有 2 幅图像不匹配语义。

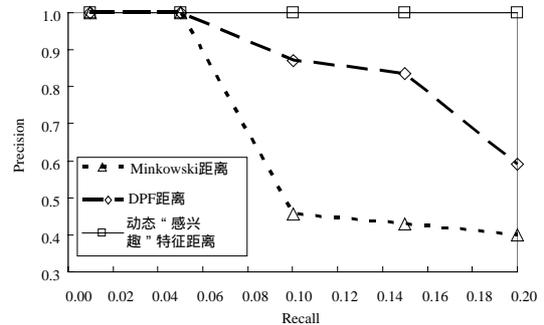
图 3 是从图像库的 4 个大类中各任意选择 5 幅图像并分别以 Minkowski 距离，DPF 距离和动态“感兴趣”特征距离检索的 Precision-Recall 比较图，每个点都是由 5 幅图像检索结果得到的 Precision-Recall 的平均值。



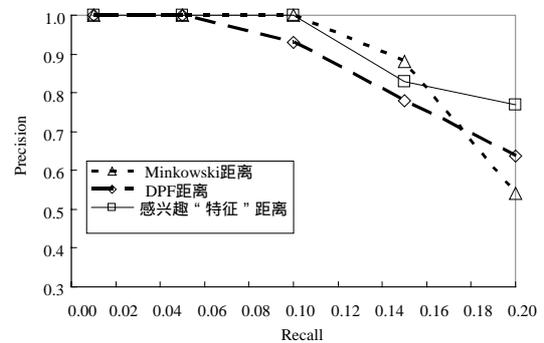
(a) 玫瑰类



(b) 雪山类



(c) 落日类



(d) 多年生植物类

图 3 不同分类的各种距离度量下的 Precision-Recall 比较

可以看到，在图 3(b)、图 3(c)中，应用“感兴趣”特征得到的检索结果最好；在图 3(a)、图 3(d)中，虽然在 Recall

(下转第 232 页)