

中图法分类号: TP391 文献标志码: A 文章编号: 1006-8961(2011)10-1850-08

论文索引信息: 支力佳, 张少敏, 赵大哲, 赵宏, 林树宽. 互信息启发的相似度组合图像检索算法[J]. 中国图象图形学报, 2011, 16(10): 1850-1857

互信息启发的相似度组合图像检索算法

支力佳^{1),2),3)}, 张少敏^{1),2),3)}, 赵大哲^{1),2),3)}, 赵宏^{1),2),3)}, 林树宽^{1),2)}

¹⁾ (东北大学信息科学与工程学院, 沈阳 110004)

²⁾ (东北大学医学影像计算教育部重点实验室, 沈阳 110004)

³⁾ (国家数字化医学影像设备工程技术研究中心, 沈阳 110004)

摘要: 图像的视觉特征与用户描述之间的差距一直是影响基于内容的图像检索准确度的最主要因素。对多种相似度进行组合来检索图像是近几年图像检索领域涌现出的一个研究热点,也是缩小这种差距的一种有效途径。如何选择更好的组合方法则是该领域很多研究者关注的核心问题。提出一种新的相似度组合算法。该算法基于互信息度量相对熵的原理,计算连续变量相似度与离散变量相似性之间的相关性,对多种相似度进行选择,以“和规则”组合相似度。在公用数据集上进行检索实验,该算法优于当前其他的“和规则”下的组合方法。

关键词: 基于内容的图像检索; 相似度组合; 互信息; 和规则

Combining similarity measures in content-based image retrieval guided by mutual information

Zhi Lijia^{1),2),3)}, Zhang Shaomin^{1),2),3)}, Zhao Dazhe^{1),2),3)}, Zhao Hong^{1),2),3)}, Lin Shukuan^{1),2)}

¹⁾ (College of Information Science and Engineering, Northeastern University, Shenyang 110004 China)

²⁾ (Key Laboratory of Medical Image Computing (Northeastern University), Ministry of Education, Shenyang 110004 China)

³⁾ (National Engineering Research Center of Digital Medical Imaging Equipment, Shenyang 110004 China)

Abstract: The lack of accordance between the information that one can extract from an image and the interpretability of the same image in a given situation is the most important factor that hampers the accuracy of content-based image retrieval (CBIR). Recently, the combination of several similarity measures draws much interest in the CBIR area, It can be shown that is effective in reducing this discordance. The core problem is: how to choose a better way to combine these similarities? In this paper, we propose a new combination algorithm. It combines similarity measures under the sum rule based on mutual information which estimates the correlation between the continuous random variable similarity measures and the discrete random variable similarity. The experimental results show that this algorithm achieves a high accuracy and efficiency in real-world image collections.

Keywords: content-based image retrieval; similarity measures combination; mutual information; sum rule

收稿日期: 2010-10-14; 修回日期: 2010-11-14

基金项目: 国家自然科学基金项目(60873009)。

第一作者简介: 支力佳(1977—), 男。东北大学计算机应用技术专业博士研究生, 主要研究方向为医学图像配准、图像检索的研究。E-mail: zhilj@neusoft.com。

0 引言

基于内容的图像检索 (CBIR) 包括各种旨在通过视觉内容组织数字图像文件的技术^[1]。CBIR 系统的目的就是 from 图像库中检索到与用户提交的查询样图在内容上一致或相似的图像集合^[1-3]。影响 CBIR 系统性能的最主要因素是语义鸿沟 (semantic gap)^[1-2]。语义鸿沟产生于对视觉图像提取出来的描述信息与用户的描述信息之间的差距。

为了提高系统性能, CBIR 领域涌现出了一些核心技术: 视觉描述、相似性测度、分类与聚类以及搜索策略^[1]。其中, 视觉描述的主要方式是通过特征向量来表达图像内容, 相似度是使用特定的视觉特征与相似性测度对两幅图像进行匹配得到的值, 反映它们之间的相似程度。如何根据不同的视觉特征构造相似性测度以及对相似度进行合理的调节是改善语义鸿沟的有效途径^[1-2]。不同的特征和相似性测度从不同方面刻画图像内容, 产生不同的相似度。在传统信息检索领域中的元搜索 (meta search)^[4] 和模式识别领域中的分类器组合^[5] 获得了良好的效果。由于图像内容本身的复杂性, 图像底层特征与高层语义之间存在很大差别, 如何对相似度进行组合, 来提高检索算法的性能也成为了 CBIR 领域近期涌现的一个研究热点^[1-2, 6-8]。Miguel 等人使用贝叶斯学习对相似度进行乘积规则 (product rule) 的组合。Torres 等人以及 Iqbal 等人在和规则下 (sum rule) 研究了相似度的组合。其中, Torres 小组使用遗传算法 (GP/GA) 对相似度的权值进行调节^[7], Iqbal 通过对相似度进行高斯标准化后使用均值权重进行组合^[8]。这些研究显示了相似度组合能够提高图像检索的准确度。然而, 上述研究中考虑的是相似度较少的情况, 为 3~5 种, 且均为使用实验者根据经验事先细致筛选过的视觉特征及相似性测度。这样会对 CBIR 性能造成影响: 1) 检索使用的相似度对先验依赖程度很高, 要求使用者提供针对被检图像集效果很好的特征与相似性测度, 对于专业用户与专门图像集较为适用, 降低了 CBIR 对于一般用户及普通图像的适用性; 2) 随着视觉特征提取技术的不断发展, 具备不同特性的视觉特征不断涌现, 过少的特征降低了 CBIR 描述图像内容的能力, 限制了检索准确度的进一步提高。

近来, 基于尺度空间关键点的局部不变特征显

示了其在刻画图像内容方面的稳定性与区分能力^[9-10]。将全局特征与局部特征结合使用, 在图像分析研究中体现了良好的性能, 也成为图像标注、检索领域的一个研究热点^[6, 11]。本文针对全局特征与局部特征的种类较多, 相似度数量较大, 用户对一般数据集的先验不足的特点, 以及检索应用环境要求响应速度快度的要求, 提出一种基于互信息 (MI) 的组合相似度的图像检索算法。该算法在概率的基础上度量连续随机变量相似度与离散随机变量相似性之间的相关性, 在贝叶斯准则下通过直接对训练数据进行采样来估计条件概率与边缘概率, 进而选取与相似性的统计相关性较高的相似度进行组合。互信息能够对相似度进行有效的筛选, 而且计算速度快, 训练抽样率低, 训练时间短。所以, 基于互信息的相似度组合方法能更好地结合多种相似度, 提高 CBIR 的适用性, 且更易于在图像检索的实际环境中使用。通过在公共数据集上与常用的相似度组合方法进行比较, 结果显示本文算法在很少训练用时的情况下, 获得了更优的检索准确度。

1 基于互信息的相似度组合图像检索算法

1.1 互信息

互信息是度量随机变量之间的相关性的一种有效工具。互信息通过一个已知随机变量的先验来降低另一个随机变量的不确定性^[12-14], 其定义为

$$I(p; q) = H(p) - H(p|q) = \sum_{x,y} r(x,y) \log_2 \frac{r(x,y)}{p(x)q(y)} \quad (1)$$

式中, $r(x, y)$ 是随机变量 x 和 y 的联合分布, $p(x)$ 和 $q(y)$ 分别是随机变量 x 和 y 的边缘分布。互信息反映了联合分布 $r(x, y)$ 与边缘分布乘积 $p(x)q(y)$ 的相对熵, 度量了变量的分布与独立统计之间的差距。在假定两个变量独立情况下, 可用边缘分布的乘积 $p(x)q(y)$ 来代表联合分布, x 和 y 的分布越相似, $p(x)$ 与 $q(y)$ 越接近, $r(x, y)$ 也越接近于 $p(x)$ (或 $q(y)$)。在极端情况下: 1) $p(x) = q(y)$, 则 x 和 y 间的互信息达到最大值 $H(p)$; 2) 当 x 和 y 完全独立时, 有 $r(x, y) = p(x)q(y)$, 此时 $I(p; q) = 0$ 。

1.2 全局特征、局部特征以及相似性测度

CBIR 中, 常见的特征基本可以分为全局特征和局部特征^[15]。全局特征是在整幅图像范围内进行

计算, CBIR 中常见的全局特征有颜色直方图、灰度共生矩阵、Tamura 纹理、Gabor 纹理等。随着尺度空间关键点提取方法的逐渐稳定^[9-10], 基于尺度空间关键点的局部特征也应用到了 CBIR 中, 并且在一些实验中取得了良好的效果^[13]。常见的尺度空间关键点有 DoG 点、Harris 点等, 而 Dense 点(即呈网格状均匀分布于图像上)也是描述图像局部特征时的常用点。而基于关键点的局部特征有 SIFT、local jet 等^[9-10]。特征向量是图像检索中最普遍使用的特征组织形式, 欧几里得距离则是在向量空间(vector space)中广泛使用的相似性测度^[1-2], 简森-仙农散度(JSD)对于特征向量为直方图的情况也是一种常用的相似性测度。

本文采用的全局特征为颜色直方图、灰度共生矩阵、Tamura 纹理和 Gabor 纹理(包括向量与直方图两种); 采用的局部特征为 DoG-sift、DoG-lj(local jet)、Harris-sift、Harris-lj、Dense10-sift 和 Dense10-lj(注, Dense10 为以 10 个像素间隔的网格点); 采用的相似性测度为欧几里得距离、简森-仙农散度以及在颜色直方图上计算的直方图交集(color histogram intersection)^[9-10, 15](注, 实际使用直方图交集相似性测度时, 对计算得到的相似度 d 作线性变换 $\hat{d} = 1 - d$, 这样就与欧几里得距离、简森-仙农散度相符, 最相似的图像相似度值接近于 0)。本实验使用的 19 种特征与相似性测度列于表 1。

表 1 特征、相似性测度及相似度表示符号

Tab. 1 Features, similarities and their representation

序号	特征	特征维数	相似性测度	相似度表示符号	
1	颜色直方图	512	欧氏距离	ColorHistEuc	
2	颜色直方图	512	简森-仙农散度	ColorHistJSD	
3	颜色直方图	512	颜色直方图交集	ColorHistInter	
4	全局特征	Gabor 向量	24	欧氏距离	GaborVecEuc
5	Gabor 直方图	96	简森-仙农散度	GaborHistJSD	
6	Tamura 特征	192	简森-仙农散度	TamuraFeatureJSD	
7	灰度共生矩阵	32	欧氏距离	CoTextureFeatureEuc	
8	DoG-sift	2 000	欧氏距离	dog. sift. histEuc	
9	DoG-sift	2 000	简森-仙农散度	dog. sift. histJSD	
10	DoG-local jet	2 000	欧氏距离	dog. lj. histEuc	
11	DoG-local jet	2 000	简森-仙农散度	dog. lj. histJSD	
12	Harris-sift	2 000	欧氏距离	har. sift. histEuc	
13	局部特征	Harris-sift	2 000	简森-仙农散度	har. sift. histJSD
14	Harris-local jet	2 000	欧氏距离	har. lj. histEuc	
15	Harris-local jet	2 000	简森-仙农散度	har. lj. histJSD	
16	Dense10-sift	2 000	欧氏距离	dense10. sift. histEuc	
17	Dense10-sift	2 000	简森-仙农散度	dense10. sift. histJSD	
18	Dense10-local jet	2 000	欧氏距离	dense10. lj. histEuc	
19	Dense10-local jet	2 000	简森-仙农散度	dense10. lj. histJSD	

对于局部特征, 采用目前较为广泛应用的视觉词直方图(visual words histogram)形式^[10, 15]。本文实验先使用 k 均值算法将局部特征聚类成 2 000 长度的码书(code book), 即聚类中心数目 $k = 2 000$ 。之后, 使用得到的码书对每幅图像进行编码, 将图像中的每一个局部特征使用与其欧几里得距离最近的

聚类中心代替, 并进行统计, 这样每幅图像就可得到一个长度为 2 000 的直方图, 该直方图即可作为这幅图像的特征。

1.3 基于互信息的相似度组合

1.3.1 组织相似度

与常见的单个特征不同, 相似度是在检索过程

中两幅图像进行匹配时产生的,相似度组合则是在这样的数据上进行的^[6-8]。多种特征的相似度组合成一个最终的相似度,用来对检索结果进行排序。图 1 表示了相似度的组合过程。

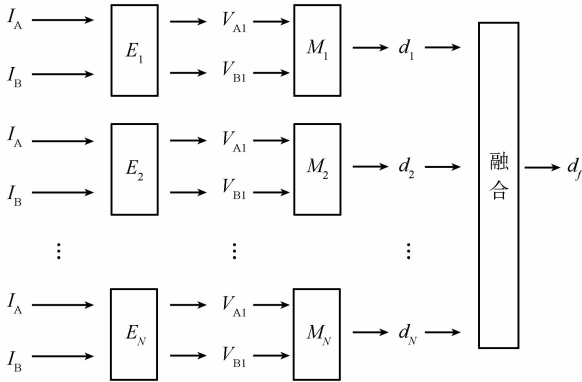


图 1 相似度的组合过程
Fig 1 Combining similarities

其中, I_A 和 I_B 分别为两幅正在匹配的图像, $E_k (k=1, \dots, N)$ 为特征提取算法, V_{Ak} 和 $V_{Bk} (k=1, \dots, N)$ 分别为提取得到的特征向量。 $M_k (k=1, \dots, N)$ 为不同的相似性测度, $d_k (k=1, \dots, N)$ 为得到的相似度值, d_j 为最后得到的组合后的相似度值。其中, 组合方法基本分为乘积规则 (product rule) 与和规则 (sum rule)^[5-7]。由于文献[7-8]中均值加权取得了很好的效果, 本文采用和规则中的均值加权 (mean rule)。

1.3.2 计算互信息

将不同的相似度看做随机变量, 可应用互信息来度量相似度与相似性的关系。

$$I(D_j, S) = H(S) - H(S | D_j) \quad (2)$$

式中, D_j 代表第 j 种相似度 (本文实验中 $j = 1, \dots, 19$), 是连续随机变量, S 代表相似与否, 是离散随机变量, 二元取值 $S = \{ \text{similar}, \text{dissimilar} \}$ 。通常假设相似与否的先验概率相等, 即 $P(\text{similar}) = P(\text{dissimilar}) = 0.5$ 。则熵为

$$\begin{aligned} H(S) &= - (P(\text{similar}) \log_2 P(\text{similar}) + \\ &P(\text{dissimilar}) \log_2 P(\text{dissimilar})) = \\ &- 2(0.5 \log_2 0.5) = 1 \end{aligned} \quad (3)$$

相对熵 $H(S | D_j)$ 同时包含离散变量和连续变量, 即

$$\begin{aligned} H(S | D_j) &= \\ &- \int p(d_j) (p(\text{similar} | d_j) \log_2 p(\text{similar} | d_j) + \end{aligned}$$

$$p(\text{dissimilar} | d_j) \log_2 p(\text{dissimilar} | d_j)) dd_j \quad (4)$$

条件概率密度 $p(\text{similar} | d_j)$ 的估计比较困难, 而 $p(\text{dissimilar} | d_j) = 1 - p(\text{similar} | d_j)$, 根据贝叶斯准则, 有

$$p(\text{similar} | d_j) = \frac{p(d_j | \text{similar}) P(\text{similar})}{p(d_j)} \quad (5)$$

需要估计条件概率密度 $p(d_j | \text{similar})$ 与边缘概率密度 $p(d_j)$ 。其中 $p(d_j)$ 可通过全概率公式进行估计, 即 $\hat{p}(d_j) = P(\text{similar}) \hat{p}(d_j | \text{similar}) + P(\text{dissimilar}) \hat{p}(d_j | \text{dissimilar})$ 也可直接对全体测度数据进行采样, 再通过 Parzen 窗进行估计^[12]。由于采样估计概率过程具有误差, 本文选择直接对相似度数据进行采样来分别估计 $p(d_j | \text{similar})$ 与 $p(d_j)$, 以减小累积误差, 即

$$\hat{p}(d_j | \text{similar}) = \frac{1}{n_{\text{similar}}} \sum_{i=1}^{n_{\text{similar}}} \frac{1}{h_n} \varphi\left(\frac{x - x_i}{h_n}\right) \quad (6)$$

式中, n_{similar} 为在训练数据中相似图像对的相似度数据上的采样数, $\varphi(\cdot)$ 为窗宽函数, h_n 为窗宽参数。其中, 高斯函数为常用的窗函数, 窗宽参数 h_n 的选择比较复杂, 本文依据较为广泛的经验, 对图像数据设置 $h_n = 0.4$ ^[13]。类似, 可以估计 $p(d_j)$, 只是采样数 n 为在训练数据中全部图像对的相似度数据上的采样数。进而通过在训练集上的再采样, 可将式的积分用求和平均来计算, 即

$$\begin{aligned} \hat{H}(S | D_j) &= \\ &- \sum_{i=1}^n \frac{1}{n} (\hat{p}(\text{similar} | d_{ji}) \log_2 \hat{p}(\text{similar} | d_{ji}) + \\ &\hat{p}(\text{dissimilar} | d_{ji}) \log_2 \hat{p}(\text{dissimilar} | d_{ji})) \end{aligned} \quad (7)$$

式中, n 为采样数。将式(3)及(5)~(7)计算得到的值, 回代入式(2), 得到第 j 种相似度与相似与否 (随机变量 S) 的互信息。依此, 可计算出每种相似度与随机变量 S 的互信息。

1.3.3 组合相似度

CBIR 中相似度的组合规则基本分为“乘积规则”与“和规则”^[6-8]。鉴于“和规则”在传统的信息检索以及图像检索中取得良好的效果, 且容易理解其组合过程, 本文基于“和规则”组合相似度。在“和规则”中又有“中值规则”、“均值规则”、“最大规则”等。由文献[5-8]实验, “均值规则”是“和规则”中性能稳定, 且在图像检索中效果良好的规则。本文使用截断相似度的“均值规则”, 具体步骤如下:

1) 在训练集上, 依次计算 m 种相似度与随机变

量 S 的互信息(本文实验中 $m = 19$);

2)对互信息值进行排序,按截断率 r ,选取互信息最大的 mr 种相似度(本文实验中 $r = 0.5$)。

使用“均值规则”,对 z -score(0 均值,单位方差)后的相似度进行组合,得到最终的组合后的相似度

$$d_f = \frac{1}{mr} \sum_{j \in I_s} d_j \quad (8)$$

式中, I_s 为截断后选取的相似度集合。

2 实验及讨论

本文的实验采用 CBIR 实验中常用的公共数据集 WANG^[15],是 Corel 照片集的一个子集。该数据集是从 Corel 数据集中手工选择出的 1 000 幅图像,包含 10 个类,每类 100 幅图像。WANG 数据集类似于通常的照片检索,从每类照片中抽出几幅进行检索^[15]。由于该数据集已由人工标注,对于每个检索图像,数据集中有 100 幅图像与之相关。

2.1 实验设计

1)训练集 估计相似度的概率密度需要训练数据。对于每一种相似度,每类相关的 100 幅图像产生 4 950(100 × 99/2) 个相似度值;而总共 1 000 幅图像共产生 499 500(1 000 × 999/2) 个相似度值。仿照 Miguel 的实验安排^[6],对每一种相似度,从每类相关的相似度值中随机抽取 1 024 个相似度来估计条件概率密度 $p(d_j | \text{similar})$ 。从全部的相似度值中随机抽取 5 000 个相似度来估计边缘概率密度 $p(d_j)$ (抽样率为 1%)。

2)测试集 CBIR 算法的性能评测大多是通过在多个检索用例上的平均性能来计算的^[6-8,15]。本文从每类图像中随机抽取 10 幅作为查询用例,这样总共有 100 幅查询用例。

3)评测方法 参见文献[6,15],本文通过在 100 幅查询用例上计算 MAP (mean average precision)来评价算法的准确度。计算 MAP 需要 3 个步骤。

(1)对于每个查询用例,统计其查准率-查全率值 $R \rightarrow P(R)$

$$P = \frac{\text{检索出的相关图像数}}{\text{检索出的全部图像数}} \quad (9)$$
$$R = \frac{\text{检索出的相关图像数}}{\text{全部的相关图像数}}$$

(2)对于每个查询用例 q ,计算 AP (average precision)值

$$AP(q) = \frac{1}{100} \sum_{n=1}^{100} P_q(R_n) \quad (10)$$

(3)在整个测试集 Q 上计算 MAP 值

$$MAP = \frac{1}{|Q|} \sum_{q \in Q} AP(q) \quad (11)$$

式中 $|Q| = 100$ 。

(4)对比较法

本文实验在相同的测试集上与使用“和规则”的高斯均值组合^[8]以及遗传算法(GA)线性组合^[7]方法进行比较。其中,GA 使用文献[7]参数设置:基本遗传算法(simple GA),100 个个体,30 代遗传。根据相似度数值变化范围,本文实验的编码长度为 30 比特。GA 的训练集是每类随机选择 30 幅图像,总共 300 幅图像进行训练。

2.2 实验结果与讨论

1)准确度及训练用时

3 种组合方法的 MAP 值与方差比较结果如图 2 所示。可以看到,高斯均值组合方法(Gaussian_Norm_Mean)与基于互信息的组合方法(MI_guide)优于 GA 参数(GA_parameters)。

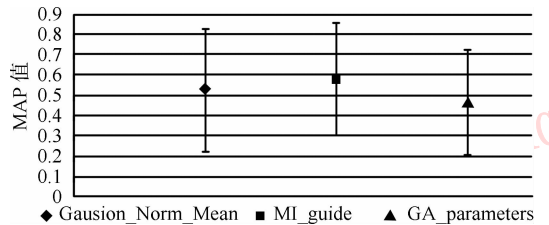


图 2 3 种组合方法的 MAP

Fig. 2 The MAP comparison of 3 combination methods

表 2 记录的 3 种组合方法的训练用时,其中高斯均值方法不需要训练。实验机器配置为 Inter (R) Core(TM)2 Quad CPU, Q6600@2.40 GHz, 2.39 GHz, 3.37 GB 内存。

表 2 3 种组合方法的训练用时

Tab. 2 The training time of 3 combination methods

	Gaussian_Norm_Mean	MI_guide	GA_parameters
训练用时/s	0	225.036	7 881.19

互信息的训练只需 3~4 min,而 GA 的训练用时则大于 2 h。图 3 为一组检索结果,每行左端的红框图像为输入的检索图像,后面图像为检索结果。



图 3 一组检索结果示例

Fig. 3 Retrieval results obtained by our proposed method

2) 准确度比较

图 4 比较了组合相似度方法与单个相似度的 MAP 值。在 3 种组合方法 Gaussian_Norm_Mean、MI_guide和 GA_parameters 中,MI_guide 的检索准确度最优,获得了最好的 MAP,且高于全部单个相似度的 MAP。Gaussian_Norm_Mean 的 MAP 低于 dense10_sift_histJSD 的 MAP, 优于其余相似度。而 GA_parameters 的 MAP 低于 dense10_sift_histJSD、dense10_lj_histJSD 与 dense10_lj_histJSD 这 3 种相似度。

图 4 还反映了对于普通图像特征与相似性测度先验不足的问题。单个相似度的 MAP 值的分布不具明显规律:同类特征不同的相似性测度,或相同的相似性测度不同的特征之间的 MAP 差距很大,难以总结出规律性的经验。对于使用 CBIR 系统的一般用户来说,很难事先准确地估计出需要使用的特征与相似性测度,因而也难以获得较好的检索准确度。各种相似度的 MAP 值、MAP 方差,以及单个相似度与相似性变量 S 的互信息值列于表 3。

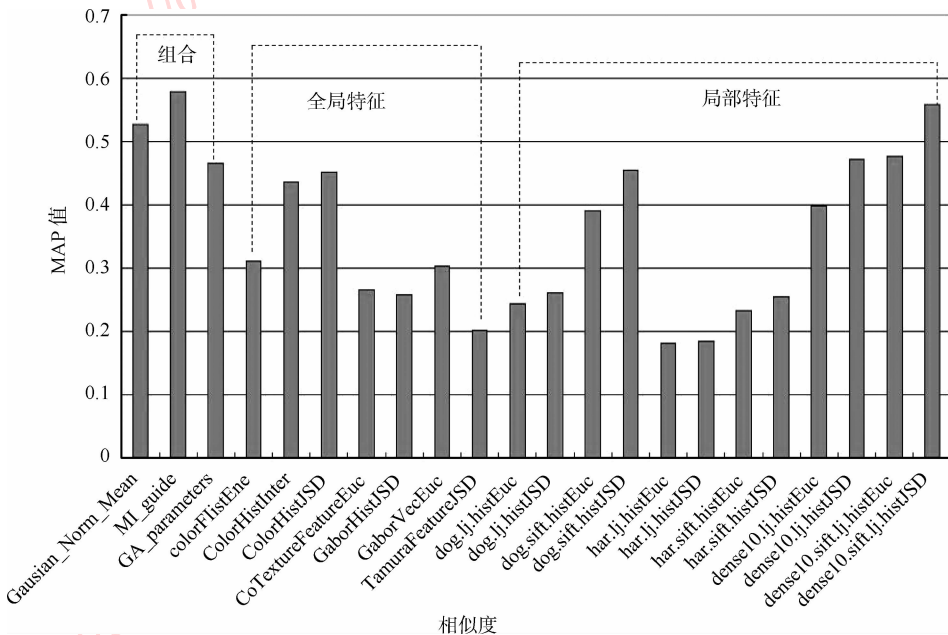


图 4 3 种组合相似度与 19 种单个相似度的 MAP

Fig. 4 The MAP comparison of 3 combination methods and 19 single similarities

表 3 组合及单个相似度的 MAP、MAP 方差与互信息
Tab. 3 MAP, MAP's stand deviation and the MI value

相似度	MAP	MAP 方差	互信息
Gaussian_Norm_Mean	0.526 8	0.303 9	
MI_guide	0.578 2	0.278 2	
GA_parameters	0.464 4	0.255 1	
colorHistEuc	0.310 7	0.237 9	0.088 8
ColorHistInter	0.435 4	0.246 8	0.079 5
ColorHistJSD	0.451 7	0.251 2	0.049 3
CoTextureFeatureEuc	0.264 6	0.255 8	0.027 1
GaborHistJSD	0.256 5	0.232 5	0.033 3
GaborVecEuc	0.303 1	0.259 7	0.044 2
TamuraFeatureJSD	0.200 7	0.176 9	0.052 2
dog. lj. histEuc	0.243 7	0.173 1	0.105 6
dog. lj. histJSD	0.261 0	0.204 4	0.007 0
dog. sift. histEuc	0.390 6	0.256 9	0.039 0
dog. sift. histJSD	0.454 8	0.297 1	0.014 3
har. lj. histEuc	0.179 9	0.154 2	0.032 2
har. lj. histJSD	0.183 6	0.162 8	0.035 0
har. sift. histEuc	0.232 9	0.180 4	0.045 5
har. sift. histJSD	0.253 6	0.212 2	0.027 3
dense10. lj. histEuc	0.398 6	0.241 5	0.304 4
dense10. lj. histJSD	0.471 6	0.279 8	0.005 8
dense10. sift. histEuc	0.475 6	0.275 3	0.026 8
dense10. sift. histJSD	0.557 9	0.297 8	0.051 0

3) 互信息与 MAP 值的符合情况

堆积折线图可显示每一数值所占大小随时间或有序类别而变化的趋势,图 5 是在测试集上的各相似度的 MAP 值与互信息值的堆积折线图。可见,由于采样误差的影响,造成了个别相似度的互信息值与在测试集上的 MAP 值差距较大(图 5 中影响选择的为 2 种相似度),但是大多数相似度与 MAP 的分布是吻合的,证明在 CBIR 中互信息可以有效地用于组合相似度,来获得更好的准确性。

3 结 论

由实验数据可见,对于 CBIR 的一般用户与普通图像,在图像检索中不同的特征与相似性测度的先验是不明确的,特征提取与相似性测度都会影响检索准确度。而且,随着特征以及相似性测度的不断增多,影响检索准确度的不确定程度也在增加。实验结果表明,高斯均值组合方法(Gaussian_Norm_Mean)对于多相似度具有较好的鲁棒性,但准确度不能超过全部单个相似度;而基于 GA 学习的组合方法(GA_parameters)则容易受到不同相似度值分布的影响,且 GA 训练时间较长;基于互信息的组合方法(MI_guide)通过对多种相似度进行合理的组合,提高了算法对于一般用户及普通图像的适用性与检索准确度,而很短的训练用时使算法更符合 CBIR 实际的应用需求(如在反馈过程中使用)。均值加权的组合方法性能稳定,如何以少量的时间代价获取更好的检索准确度则是下一步的目标。

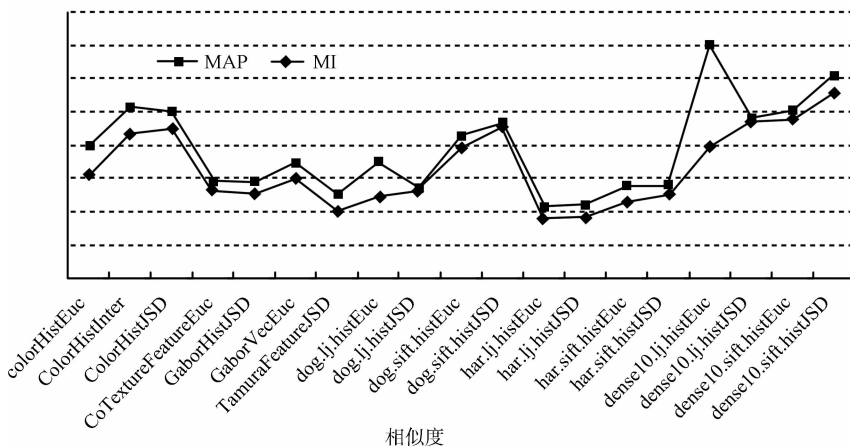


图 5 通过训练集计算的各相似度与测试集上用例的 MAP 符合程度

Fig. 5 The accordance between the similarities selected from the training set and the MAP get from the test set

参考文献 (References)

- [1] Datta R, Joshi D, Li J, et al. Image retrieval: ideas, influences, and trends of the new age[J]. *ACM Transactions on Computing Surveys*, 2008, 40(2): 1-60.
- [2] Smeulders A W, Worring M, Santini S, et al. Content-based image retrieval at the end of the early years[J]. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2000, 22(12): 1349-1380.
- [3] Su Zhong, Zhang Hong jiang, Ma Shaoping. An image retrieval relevance feedback algorithm based on the Bayesian classifier [J]. *Journal of Software*, 2002, 13(10): 2001-2006. [苏中, 张宏江, 马少平. 基于贝叶斯分类器的图像检索相关反馈算法[J]. *软件学报*, 2002, 13(10): 2001-2006.]
- [4] Elena Renda M, Umberto Strccia. Web metasearch: rank vs. score based rank aggregation methods[C] //ACM Symposium on Applied Computing. Melbourne, Florida: ACM, 2003: 841-846.
- [5] Kittler J, Hatel M, Duin R P W, et al. On combining classifiers [J]. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1998, 20(3): 226-239.
- [6] Arevalillo-Herrez M, Domingo J, Ferri F J. Combining similarity measures in content-based image retrieval [J]. *Pattern Recognition Letters*, 2008, 29(16): 2174-2181.
- [7] Torres R S, Falcao A X, Goncalves M A, et al. A genetic programming framework for content-based image retrieval[J]. *Pattern Recognition*, 2009, 42(2): 283-292.
- [8] Iqbal Q, Aggarwal J. Combining structure, color and texture for image retrieval; a performance evaluation [C] //Proceedings of 16th International Conference on Pattern Recognition (ICPR). Quebec City: IEEE, 2002:438-443.
- [9] David G Lowe. Distinctive image features from scale-invariance keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [10] Mikolajczyk K, Schmid C. Scale & affine invariant interest point detectors[J]. *International Journal of Computer Vision*, 2004, 60(1): 63-86.
- [11] Zhang Q, Izquierdo E. Optimizing metrics combining low-level visual descriptors for image annotation and retrieval [C] // Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). Toulouse: IEEE, 2006:405-408.
- [12] Duda R O, Hart P E, Stork D G. *Pattern Classification*[M]. 2nd ed. Beijing: China Machine Press, 2003: 630-633.
- [13] Peng H C, Long F H, Chris Ding. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy [J]. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2005, 27(8): 1226-1238.
- [14] Kwak N, Choi C H. Input feature selection by mutual information based on Parzen window[J]. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2002, 24(12): 1667-1671.
- [15] Thomas D, Daniel K, Hermann H. Features for image retrieval: an experimental comparison[J]. *Inf. Retrieval*, 2008, 11(2): 77-107.