

图像检索中一种新的相关反馈机制^{*}

A Novel Relevance Feedback Technique for Interactive Content-Based Image Retrieval

曹 奎 冯玉才 王元珍

(华中科技大学计算机学院数据库与多媒体研究所 武汉430074)

Abstract Content-based image retrieval systems require the development of relevance feedback mechanism to facilitate the human-computer interaction. In this paper, we propose a grey approach to relevance feedback, where the GRA is used to describe the feature distributions of the images judged relevant by the user and dynamically updates both the similarity measure and the query in order to accurately represent the user's particular information needs. Experimental results show that the proposed approach captures the user's information needs more precisely, and improves the effectiveness of image retrieval considerably.

Keywords Content-based image retrieval (CBIR), Image representation and retrieval, Relevance feedback, Grey relational analysis (GRA)

1. 引言

近年来,随着图像数据的大量增加,对图像进行智能化、基于内容的处理愈来愈引起人们的极大兴趣。除了使用手工输入的关键字外(基于文本的方法),其于内容的图像检索(Content-Based Image Retrieval, CBIR)系统也使用图像的视觉特征来建立图像索引,所使用的图像低层特征包括:颜色、纹理和形状等。在CBIR系统中,最常用的检索模式是“例子查询”(Query By Example, QBE),即首先由用户选取一个或多个例子图像(Query Images),然后,系统自动地从图像库中检索出与其相似的数据库图像。在这种检索模式中,系统的检索能力完全依赖于图像特征的性质以及用户所选取的例子图像的质量。到目前为止,大多数研究均侧重于寻找一种最好的图像表示方法,而对于人在CBIR系统中的作用以及检索过程中的相关反馈技术的研究则相当粗略。例如,典型的QBIC系统提供的交互能力较弱,它要求用户选取检索所需的图像低层特征^[1]。另外一些系统为用户提供了一些不直观的加权机制,要求用户指定不同特征在检索中的作用。所有这些均需要用户对图像特征有较深入的了解,从而增加了用户的认知负担。

考察现有的商业或研究型的CBIR系统^[1,2],虽然这些系统提供了丰富的图像表示方法,然而大多数系统却显得“僵硬”,即对于给定的例子图像和检索特征,系统均以一个固定的图像集作为检索结果。用户真正需要的是一个“柔性”系统,它能适应用户的查询需求,从数据库中检索出更多用户感兴趣的图像。这种技术在传统信息检索(IR)领域被称为“相关反馈”,它是基于文本的信息检索系统中的有力工具。所谓“相关反馈”就是通过人机交互方式在检索结果中标记或选定用户认为相似的或不相似的图像,系统利用用户提供的反馈信息自动调整当前查询,使之能较好地描述用户的查询要求。如此不断重复进行,直到用户获得满意结果为止。在CBIR领域, Picard 和 Minka 首先对相关反馈技术进行了研究^[3,4], Cox 等人提出了在 Bayesian 框架中处理相关反馈^[5],并在其中明确地加入了一个用户选择模型。文^[6,7]中也提出了一些其它的相关反馈机制。一般而言, CBIR 中的相关反馈技术可分为两类:一类是利用用户的反馈信息调整相似度量中的一些参数^[7];另一类则是从概率的观点出发,利用用户的反馈信息,

计算出图像库中每一幅图像符合用户需求的概率,并将概率高的图像作为检索结果返回用户^[8]。

自从“灰色系统理论”提出以来^[9], 灰关联分析(GRA)方法已应用于许多领域。近年来, GRA 在诸如图像编码、模式识别和聚类分析等领域的研究相当活跃^[9,10]。一般地, GRA 是一种描述给定离散灰空间中主因素与其它因素之间关系的方法。在研究具有不确定性和不完全性系统中,由于没有物理原型,信息难以完全判断,而且数据很少(即小样本),传统的统计方法很难适用,而 GRA 方法却可以获得满意的结果。由于CBIR研究的问题正属于这种情况,因此, GRA 理论就可以作为一种研究“相关反馈”的有力工具。

本文提出一种新的基于 GRA 的相关反馈技术,它使用 GRA 来描述“例子图像”与“相关图像”之间的关系,据此自动更新图像的相似度量和用户的查询需求。

2. CBIR 系统中不同特征的组合

在讨论我们提出的相关反馈技术之前,首先讨论图像的模式化方法^[6]。

设 I 为任一幅图像,则图像 I 可以用一组从图像中自动抽取出的低层特征来表示(图像特征包括颜色、纹理和形状等),即

$$I = I(D, F, R) \quad (1)$$

其中, D 为图像数据, $F = \{f_i\}$ 为图像 I 的视觉特征集合, $R = \{r_{ij}\}$, $i = 1, 2, \dots, p$, $j = 1, 2, \dots, q$, 为对应于特征 f_i 的 q 种图像表示。例如,颜色直方图(Color Histogram)、颜色矩(Color Moments)以及颜色集(Color Sets)是图像颜色特征的三种表示方法。

图像表示 r_{ij} 可视为特征空间中的一个特征向量,即

$$r_{ij} = (r_{ij}(1), r_{ij}(2), \dots, r_{ij}(K)) \quad (2)$$

这里, K 为特征向量 r_{ij} 的长度。

这样,一个CBIR模型可表示为: (D, F, R, M) , 它是由图像对象模型 $I(D, F, R)$ 和一组相似度量或距离度量 $M = \{d_{ij}\}$ 确定的。图像的距离度量 d_{ij} 用于计算两个图像之间的相似度。在实际的检索系统中,不同的图像表示可能采用不同的距离度量。例如,颜色直方图一般使用“直方图相交算子”,而“颜色矩”表示常采用 L_2 距离。

在我们给出的检索模型中,查询图像 Q 与数据库图像 I

^{*} 本文受国家863高科技项目基金资助(项目编号:863-511-920-001)。

之间的全局距离定义为:

$$d(Q, I) = \sum_{i=1}^p \sum_{j=1}^{q_i} w_{ij} d_{ij}(r_{ij}, \bar{r}_{ij}) \quad (3)$$

这里, $Q=Q(D, F, R)$, $I=I(\bar{D}, \bar{F}, \bar{R})$, w_{ij} 为图像表示 r_{ij} 在全局距离计算中的权重。

由于全局距离度量式(3)是不同图像表示间距离的线性组合, 为了使式(3)中各个距离之间具有可比性, 就需要对其进行归一化处理, 我们采用 Gaussian 归一化方法^[6]。这里, 我们考虑两种情况, 其一是将子特征距离归一化为同一取值区间, 使这些距离在式(3)中的作用相同; 另一种是对特征向量之分量进行归一化处理。

3. 相关反馈机制

如前所述, 大多数现有的 CBIR 系统所提供的交互能力都较弱(如 QBIC 系统)。有些系统为用户提供了一个非直观的特征加权机制, 允许用户指定不同特征的权重。这种固定加权机制的缺点之一是不具有一般性, 更重要的是它不能有效地模型化“高层概念”以及人的视觉感知的主观性。

(1) 对于每一种图像表示 r_{ij} ($i=1, 2, \dots, p; j=1, 2, \dots, q_i$), 做(2)~(3)。

(2) 将 r_{ij}^0 作为参考数列, $\{r_{ij}^1, r_{ij}^2, \dots, r_{ij}^m\}$ 作为比较数列(设这些数列的长度均为 q), 进行如下灰关联分析。

(2.1) 数列初始化(预处理)

$$\hat{r}_{ij}^l(l) = \begin{cases} r_{ij}^l(l)/r_{ij}^0(l) & \text{if } r_{ij}^l(l) < r_{ij}^0(l) \\ r_{ij}^l(l)/r_{ij}^0(l) & \text{otherwise} \end{cases} \quad k=0, 1, \dots, m; l=1, 2, \dots, q \quad (4)$$

(2.2) 计算灰关联系数

$$\xi_{ij}^l(l) = \frac{\min_{l'} |\hat{r}_{ij}^0(l) - \hat{r}_{ij}^l(l)| + \zeta \max_{l'} |\hat{r}_{ij}^0(l) - \hat{r}_{ij}^l(l)|}{|\hat{r}_{ij}^0(l) - \hat{r}_{ij}^l(l)| + \zeta \max_{l'} |\hat{r}_{ij}^0(l) - \hat{r}_{ij}^l(l)|} \quad (5)$$

其中, $\zeta \in (0, 1)$ 为分辨系数, 在该算法中取 $\zeta = 0.5$ 。

(3) 生成新的查询向量 $\bar{r}_{ij}^1$: 对于用户标记的那些相关图像, 将它们相应的灰关联系数 $\xi_{ij}^l(l)$ 组成一个 $m \times q$ 矩阵, 即

$$\xi = \begin{bmatrix} \xi_{ij}^1(1) & \xi_{ij}^1(2) & \dots & \xi_{ij}^1(q) \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{ij}^m(1) & \xi_{ij}^m(2) & \dots & \xi_{ij}^m(q) \end{bmatrix} \quad (6)$$

从直观上看, 式(5)计算出的这些灰关联系数反映了相应的相关图像对用户查询需求的相对重要性。灰关联系数越大, 其相对重要性越高(即越能代表用户的查询需求), 反之亦然。于是, 新查询向量的计算如下:

$$\bar{r}_{ij}^1(l) = \sum_{k=1}^m [\xi(k, l) r_{ij}^k(l) / \sum_{k=1}^m \xi(k, l)] \quad l=1, 2, \dots, q \quad (7)$$

(4) 更新权重: 对于图像表示 r_{ij} , 分别计算相关集 R 中每个图像与例子图像在该子特征上的距离, 并且将这 m 个距离组成数列, 记为 s_{ij} , $i=1, 2, \dots, p; j=1, 2, \dots, q_i$ 。设 $s_0 = (0, 0, \dots, 0)$ 为参考数列, s_{ij} 作为比较数列, 进行如下灰关联分析。

(4.1) 计算灰关联系数

$$\xi_{ij}(l) = \frac{\min_{l'} |s_0(l) - s_{ij}(l)| + \zeta \max_{l'} |s_0(l) - s_{ij}(l)|}{|s_0(l) - s_{ij}(l)| + \zeta \max_{l'} |s_0(l) - s_{ij}(l)|} \quad (8)$$

其中, $\zeta \in (0, 1)$ 为分辨系数, 在该算法中取 $\zeta = 0.5$ 。

(4.2) 计算灰关联度

$$\bar{\gamma}_{ij} = \frac{1}{m} \sum_{l=1}^m \xi_{ij}(l), \quad i=1, 2, \dots, p; j=1, 2, \dots, q_i \quad (9)$$

灰关联度反映了某一图像表示 r_{ij} 在用户进行相关图像的判断中所起的作用。 $\bar{\gamma}_{ij}$ 的值越大, 其相对重要性越大, 反之亦然。

(4.3) 计算新权重 $\bar{w}_{ij}$

$$\bar{w}_{ij} = \bar{\gamma}_{ij} / \sum_{i=1}^p \sum_{j=1}^{q_i} \bar{\gamma}_{ij} \quad (10)$$

上述算法的主要特点是: 它允许用户提交一个较粗略的初始查询, 然后, 通过相关反馈不断地求精查询, 直到满足用户的查询要求为止。该算法免除了需要用户理解图像表示和相似度量的负担, 用户仅需依据自己的视觉感知做出相关判断即可。根据用户在检索结果中所作的相关选择, 该算法能动态调整查询向量、更新权重, 以使查询能更准确地表示用户的检索要求。与 Bayesian 相关反馈技术相比^[5], 该算法的显著优点是我们无须对算法的正确性作任何假设(例如特征向量的分量相互独立), 并且算法的计算复杂度较低。

3.2 图像检索的处理过程

为了解决这些问题, 可以让用户参与检索过程, 并且根据用户的反馈信息动态地更新查询以反映用户的查询需求。我们采用 GRA 来计算相关图像的特征分布, 进而推断出用户在进行相似判断时考虑了何种特征, 以及这些特征在用户判断中的重要程度如何, 在此基础上就可以更新用户查询, 进行下一次人机交互。

3.1 详细算法

我们的相关反馈模型是: 在图像检索过程中, 允许用户在检索结果集中选取他认为与例子图像相关的图像; 然后, 应用 GRA 来分析例子图像与这些相关图像之间的关系, 并且计算相应的“灰关联系数”和“灰关联度”; 最后, 通过组合这些“灰关联系数”就可以计算出新查询向量和更新的权重。

设 Q 为一个例子图像, $\{r_{ij}^0, i=1, 2, \dots, p, j=1, 2, \dots, q_i\}$ 为图像 Q 的特性向量集。又设用户在一次检索结果中选取的相关图像集为 R , m 为 R 中图像的个数, $\{r_{ij}^k, i=1, 2, \dots, p, j=1, 2, \dots, q_i\} (k=1, 2, \dots, m)$ 为相关集 R 中图像的特征向量。相关反馈算法如下:

基于上述相关反馈机制的图像检索过程为:

(1) 用户浏览图像库, 寻找感兴趣的图像作为“例子图像”。用户也可以直接选取一个外部图像作为“例子图像”;

(2) 初始化特征的权重 $W = [w_{ij}]$, 初始设定每个图像表示在检索中具有相同重要性;

(3) 应用式(3), 计算例子图像与数据库图像之间的距离, 并且将若干个最相似图像作为检索结果返回用户;

(4) 用户在检索结果中选取或标记相关图像;

(5) 应用式(7)计算新的查询向量;

(6) 应用式(10)更新距离计算中的权重;

(7)提交新查询,转(3)开始下一次循环。

4. 实验结果与讨论

在实验中,我们选取了一个由910幅真彩色图像组成的测试集,图像集中包含有自然景物、花草、日落、蝴蝶、植物以及其它包含多种对象的图像等,这些图像是从 Internet 上获得的(<http://www.stanford.edu>)。我们知道,图像检索系统的性能与所抽取出的图像特征以及相应的图像特征表示密切相关。因此,实验的目的侧重于评价采用了本文提出的“相关反馈”技术后,图像检索性能的提高程度如何。

为了便于实现,我们选取图像的颜色和纹理特征来描述图像的内容,即

$$F = \{f_i\} = \{Color, Texture\} \quad (11)$$

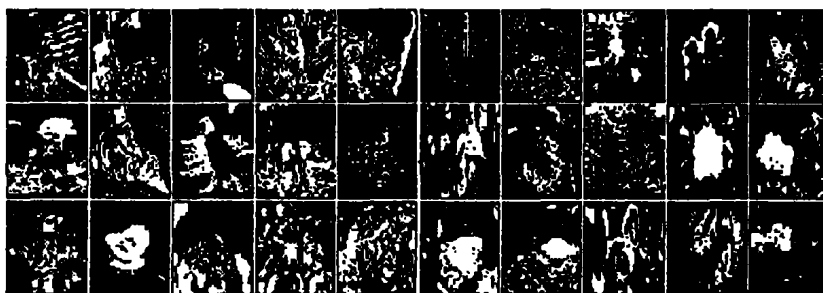


图1 初始检索结果 (0 rf)

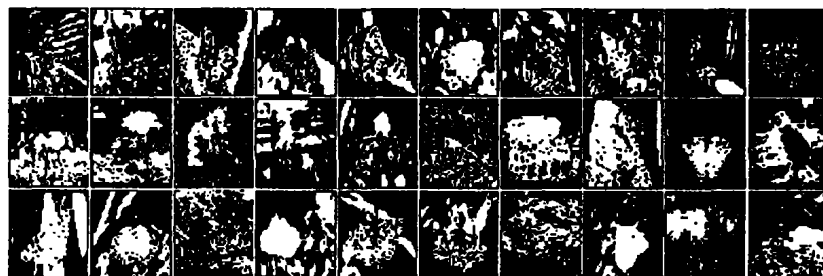


图2 第一次相关反馈后的检索结果 (1 rf)

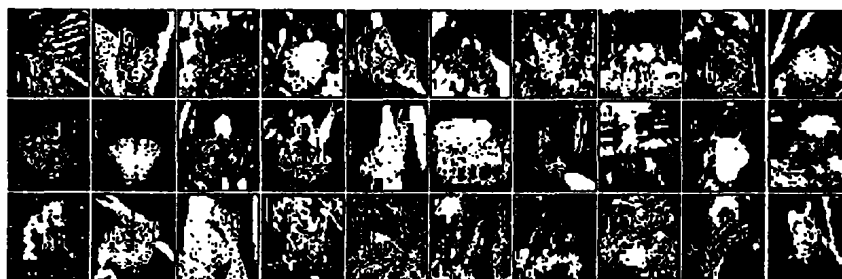


图3 第二次相关反馈后的检索结果 (2 rf)

为了较全面地评价本文提出的相关反馈技术,我们设计了5类查询:蝴蝶、花、自然景物、山、日落。图像检索性能的评价,采用一种称为“检索率”(Retrieval Rate)的量化方法来比较系统输出的检索结果与人们期望结果的一致性。检索率定义为^[12]:

$$Rate = \bar{N} / N \quad (12)$$

其中, N 表示图像库中与例子图像相似的图像个数,所有这些图像组成一个查询集合 QS ; $\bar{N}$ 表示检索结果的前 N 个图像中属于集合 QS 的图像个数。显然,为了使用这种评价指标,我们需要事先确定每个查询的相关图像集合,即集合 QS 。

这5类图像查询的实验结果如表1所示(符号“rf”表示相关反馈的次数)。表中给出了每类20个查询进行3次相关反馈的平均检索率,其中初始查询中不同图像特征赋予相同的权

采用“颜色直方图”来描述图像的颜色特征,相似度量采用 L_1 距离;图像纹理特征的表示则采用“共现矩阵”法^[11],相似度量采用 L_2 距离。

图1~3给出了检索一个“蝴蝶”图像的检索过程,图中左上角图像为“例子图像”,从上到下、自左至右按相似度的降序将前30幅图像返回用户。图1给出了该查询的初始检索结果,其中包含一些“错误图像”;在图1所显示的检索结果中,用户选取若干个相关图像,基于用户的反馈信息,系统重新计算查询向量并且更新权重,然后进行下次循环。这样,就会有更多的相关图像被检索出来(如图2、图3所示)。在检索示例中,第一次查询中各特征的权重是相同的,之后,依据用户的反馈信息,系统动态地调整特征的权重,将符合用户视觉感知的特征赋予较大的权重。

重。实验结果表明,采用本文提出的相关反馈机制能显著地改善图像检索性能(平均检索率从66.5%提高到86.8%)。

表1 5类图像检索的平均检索率

类别	Rate(0 rf)	Rate(1 rf)	Rate(2 rf)
蝴蝶	60.3%	78.2%	80.2%
花	77.6%	93.8%	96.3%
自然景物	55.2%	70.8%	75.8%
山	70.7%	87.6%	91.6%
日落	68.6%	83.7%	90.3%

结论 在CBIR研究中,有两个基本问题需要解决。其一是图像特征的表示方法;另一个是学习机制的设计。然而,大多数早期研究都侧重于寻找一种“最好”的图像表示,检索的

Ontology 建模方法研究^{*}

The Methodology of Ontology Building

徐振宁 黄凯歌 张维明 陈文伟

(国防科技大学管理科学与工程系 长沙410073)

Abstract The ontology model of a certain domain is an effective approach for the intercommunion between people from the different domains, the communication and interoperation among agents, and the share and reuse of the software. But the lack of formal analysis tools for domain modeling results in taking liberties with conceptualization. This paper discusses how to introduce the ontological notions from philosophy into Knowledge Engineering in order to supply a set of formal analysis tools for conceptualization analysis. This method can definitely record the hypothesis and analysis criterions through ontology building, and clarify the hierarchy of concepts.

Keywords Ontology, Conceptualization, Ontological analysis, Meta-property, Life cycle

1 引言

形式化表示的知识来源于对研究领域的概念化,这种概念化的主要内容包括:研究领域的对象、概念和其它假定存在的实体,以及维持它们的关系^[1]。概念化是人们出于某种目的对希望进行表示的世界的一个抽象的、简化的视图。每一个知识库,基于知识的系统,知识级的主体,甚至每一个信息系统都隐式或明确地建立在某种概念化的基础上。

本体(Ontology)是一个源于哲学的概念,是一种对“存在”的系统化的解释。在人工智能领域,本体被界定为概念化的某些方面的明确的说明或表示^[2]。当研究领域中的知识以说明的形式进行表示时,领域中能够被表示的对象的集合被称为论域。通过设计能够反映这些对象和对象之间关系的具有代表性的词汇集,对研究领域的本体进行描述。这样的本体包括领域核心的概念、关系、实例、公理等实体。论域中实体的名称通过定义,与人类能够阅读的描述实体名称意义的文本和约束这些名称如何解释和应用的公理相关联。

知识工程领域的研究者在其研究领域中引入这一概念,用于更好地解决知识共享中的问题。本体技术能够有效地促进来自不同领域的研究人员或组织间的交流,软件主体间的通信和互操作,异构数据源的集成,数字图书馆,以及加强知识系统的可重用性和系统的可靠性。因此,不同领域的研究者开发了大量领域相关的本体,本体技术得到了广泛的应用。

2 相关工作

文[3,4]总结了作者在化学领域构造本体的经验,归纳了作者在工程实践中符合本体开发过程的行为集,在进化快速原型方法基础上建立了本体开发的生命周期,提出了一种本体的结构化开发方法,以及开发过程中多种中间结果的表示技术。文[5]从工程的角度总结了形式化本体的开发过程,提出了指导这种开发过程的一组标准,以及将这些标准应用于应用领域的示例研究。

企业本体(Enterprise Ontology, EO)是大型企业建模基础设施的一个重要组成部分,它覆盖了企业建模所有的核心

^{*} 本课题得到国家自然科学基金(项目编号 79800007)和国防科技大学预研基金资助。徐振宁 博士研究生,主要研究方向为分布式人工智能,多主体系统;黄凯歌 博士研究生,主要研究方向为分布式人工智能,Web 数据开采;张维明 教授,主要研究方向为信息系统,智能决策支持系统,软件工程;陈文伟 教授,博士生导师,主要研究方向为人工智能、决策支持系统、智能决策支持技术。

实现基于不同特征相似性的组合(特征的权重一般是固定的)。本文给出了一种新的相关反馈技术,它通过从用户交互中学习,逐步求精查询和模型化用户的视觉感知,从而更准确地描述用户的查询需求。与其它基于统计方法的相关反馈机制不同,我们采用 GRA 理论来计算“例子图像”与“相关图像”之间的关系,并据此更新查询向量和特征的权重。实验结果表明,本文提出的相关反馈机制能较好地符合人的视觉感知的主观性,从而显著地改善图像检索的性能。

参考文献

- 1 Niblack W, et al. The QBIC project: querying images by content using color, texture and shape. SPIE Proc. Storage and Retrieval for Image and Video Databases, 1993. 173~187
- 2 Pentland A, Picard R, Scharoff S. Photobook: Content-based manipulation of image databases. International Journal of Computer Vision, 1996, 18(3): 233~254
- 3 Picard R, Minka T, Sdaroff S S. Modeling subjectivity in image libraries, IEEE Int. Conf. On Image process, Lausann
- 4 Minka T, Picard R. Interactive learning using a society of models. Pattern Recognition, 1997, 30(4): 565~581
- 5 Cox I, et al. Pichunter: Bayesian relevance feedback for image retrieval. In: Proc. of 13th Int. Conf. on Pattern Recognition, Vienna, Australia
- 6 Yong Rui, et al. Relevance feedback techniques in interactive content-based image retrieval. In: Proc. of IS & I and SPIE Storage and Retrieval of Image and Video Databases VI, San Jose, CA, 1998. 24~30
- 7 Ciocca G, Schettini R. A relevance feedback mechanism for content-based image retrieval. Information Processing and Management, 1999, 35: 605~632
- 8 Deng J. Control systems of grey systems. Systems and Control Letter, 1982, 5: 288~294
- 9 Hsu Y T, et al. High noise vehicle plate recognition using grey system. The Journal of Grey System, 1998, 10: 193~208
- 10 Hsieh C H. Grey image hiding. The Journal of Grey System, 2000, 3: 275~282
- 11 Furht B, et al. Video and image processing in multimedia systems. Boston: Kluwer Academic Publishers, 1995. 226~270
- 12 Androusos D, Plataniotis K N, Venetsanopoulos A N. A novel vector-based approach to color image retrieval using a vector angular-based distance measure. Computer Vision and Image Understanding, 1999, 75(1/2): 46~58