



(12) 发明专利申请

(10) 申请公布号 CN 103164539 A

(43) 申请公布日 2013. 06. 19

(21) 申请号 201310128036. 3

(22) 申请日 2013. 04. 15

(71) 申请人 中国传媒大学

地址 100024 北京市朝阳区定福庄东街 1 号

(72) 发明人 黄祥林 杨丽芳 曾南子 韩卯辉

(51) Int. Cl.

G06F 17/30 (2006. 01)

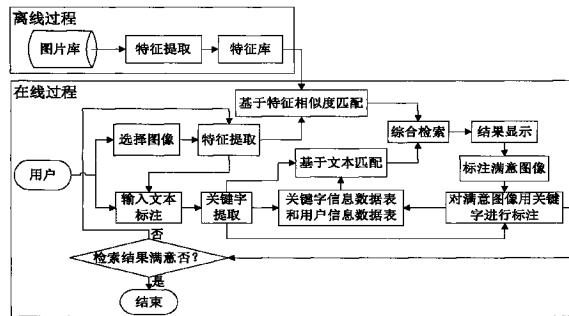
权利要求书2页 说明书5页 附图2页

(54) 发明名称

一种结合用户评价与标注的交互式图像检索方法

(57) 摘要

本发明是一种结合用户评价和标注的交互式图像检索方法，属于多媒体信息检索领域。在本发明中，该方法利用了基于图像的物理特征和文本相结合的综合检索方法，在检索过程中，允许用户对查询图像进行文本信息描述，或者选择系统提供的关键字，通过对检索结果进行“满意”或“不满意”的相关评价，图像检索系统自动对用户标记的相关满意图像进行文本标记，形成高层语义信息；随着用户的不断使用，该系统会生成丰富的语义信息数据库。考虑到不同用户对同一图片，同一用户不同时间对同一图片文本标注的差异，本发明在生成语义信息数据库的过程中结合了用户的可信度。进行检索时，对存在语义信息的查询图像采用基于特征和文本相结合的综合检索方式进行检索，提高了检索结果的准确度。本发明具有效率高、精确度高、交互方式友好的优点。



1. 一种结合用户评价和标注的交互式图像检索方法,其特征在于具体步骤如下:

1) 对图像库中的所有图像进行特征提取,离线生成图像特征库 V;

2) 用户登录,如果不登录即为匿名用户;

3) 用户提交查询信息,该信息可以是查询图像  $Q_N$ ,也可以是输入的文本信息,或者是查询图像与文本信息同时存在;

4) 检索系统对查询图像  $Q_N$  进行特征提取,并对输入的文本信息提取关键字;

5) 检索系统对查询图像  $Q_N$  进行特征相似度匹配 S1;

6) 检索系统对提取的关键字进行文本的匹配,即语义相关性 S2;

7) 检索系统利用 S1 和 S2 计算总相似度 S,返回 S 值最大的图像集作为初始的检索结果集合 set(R);

8) 同时,检索系统在关键字数据表 KL 中查询是否存在 set(R) 中各图像名对应的记录,如果存在,则给出这些记录对应的关键字权值排序最高的前 N( $N > 0$ ) 个关键字供用户选择;

9) 用户选择系统提供的关键字 K,同时也可对图像文本标注  $T'$ ;

10) 系统根据用户选择的关键字和标注的文本信息,组合新关键字  $K'$ ;

11) 同时,用户对结果图像进行“满意”或“不满意”评价,对标记为“满意”的图像,用关键字  $K'$  进行文本标注;

12) 图像检索系统计算该用户的用户可信度 a,并计算“满意”图像标注的关键字权值 w,保存到关键字信息表 KL 中,同时,将注册用户的用户信息保存到用户信息数据表 UL 中,匿名用户的用户信息不需要保存,只需对 KL 进行更新;

13) 如果用户对检索结果不满意,则可修改文本标注 T,重复步骤 (3)-(12),直到用户对检索结果满意为止,并给出最终的检索结果。

2. 根据权利要求 1、4 所述提取的图像特征,其特征在于该特征包括:颜色特征、纹理特征、形状特征、空间分布特征及局部不变性特征。

3. 根据权利要求 4 所述的关键字,其特征在于提取的关键字可以是单一的,也可以是多个关键字。

4. 根据权利要求 1 所述的图像检索方法,其特征在于:步骤 7 所述的图像库总相似度 S 的计算方法如下:

$S = \alpha S_1 + \beta S_2$ ,其中  $\alpha + \beta = 1$ ,  $\alpha$  是基于特征相似度匹配在综合检索中所占的比重,  $\beta$  是基于文本匹配在综合检索中所占的比重。

5. 根据权利要求 1 所述的图像检索方法,其特征在于:步骤 8、12 所述的关键字信息数据表 KL 的字段如下:

关键字、图像名、关键字评价次数、关键字权值。

6. 根据权利要求 1 所述的图像检索方法,其特征在于:步骤 12 中所述的用户信息数据表 UL 的字段如下:

用户名、密码、用户评价次数、用户可信度。

7. 根据权利要求 1 所述的图像检索方法,其特征在于:步骤 12 中所述的 KL 和 UL 的更新,具体规则如下:

1) KL 的更新:若某次检索过程取得的 K 中的某个关键字与用户标注为满意的某个图像

名的对应记录不存在，则将该关键字、该图像的图像名、关键字评价次数（1次）和该关键字对该图像名的关键字权值作为一条记录存入 KL 中，若记录存在，则将该条记录的评价次数加 1，并重新计算该关键字对应记录的关键字权值项；

2) UL 的更新：若进行检索的是注册用户，则将该用户对应记录的用户评价次数加 1，并重新计算该注册用户的用户可信度。

8. 根据权利要求 1 所述的图像检索方法，其特征在于：步骤 12 所述的用户可信度 a，对于匿名用户，它的可信度值只取决于本次操作；而对于注册用户，它的用户可信度是该用户多次评价的用户可信度的平均值；用户某次操作的用户可信度 a 的具体计算方法为：从 KL 中找到包含 K（用户选择的关键字以及用户输入的文本信息整合后的关键字集合）中任一关键字的所有记录，从这些记录中找到关键字权值最高的记录，a 的大小等于该条记录对应的关键字评价次数在对查询图像进行标注的所有关键字评价次数总和中所占比重对应的量化值；

1) 对于注册用户，注册用户的第 n 次操作的用户可信度的平均值为：

$$a_n = \frac{a_{n-1} * (n - 1) + a}{n};$$

2) 对于匿名用户，匿名用户的用户可信度值为 a。

9. 根据权利要求 8 所述的某次用户可信度计算时，关键字的评价次数在对查询图像进行标注的所有关键字评价次数总和中所占比重对应的量化值的量化方法可以是均匀量化，也可以是非均匀量化。

10. 根据权利要求 8 所述的关键字信息数据表 KL 中的“关键字权值”字段，某一关键字记录中的关键字权值计算方法如下：

$$\text{关键字权值} = \sum_{i=1}^m \text{对该图像采用该关键字进行第 } i \text{ 次评价的用户可信度} ,$$

m 为该关键字的关键字评价次数。

## 一种结合用户评价与标注的交互式图像检索方法

### 技术领域

[0001] 本发明是一种结合用户评价与标注的交互式图像检索方法，属于多媒体信息检索领域。

### 背景技术

[0002] 近年来，随着信息技术的发展，数字图像的数量快速增长，每天都有海量的图像数据产生。如何快速而准确地查找并有效利用这些图像成为迫切需要解决的问题，这就需要图像检索技术。

[0003] 20世纪70年代开始出现基于文本的图像检索技术，主要利用数据库管理系统，对图像进行人工标注，即对每幅图像标注关键字，利用原有成熟的文本检索系统对图像进行管理和检索。这种方法的主要优点是检索速度快，但是人工标注费时费力、主观性强，且对海量图像进行人工标注是很难实现的。

[0004] 20世纪90年代，出现了基于内容的图像检索技术，它自动提取图像的底层物理特征（如颜色、纹理、形状等特征）来进行图像相似度比较。检索时，用户选择一幅图像作为查询图像，然后通过图像底层物理特征的匹配，由系统查找与查询图像相似的图像，按相似度大小排列并返回给用户。在实际应用中，颜色、纹理、形状等底层物理特征有时不能真实反映图像内容，即存在一定的语义鸿沟。近期，人们又提出了基于语义的图像检索方法，但如何自动获取图像的语义信息却很困难。

### 发明内容

[0005] 本发明的目的在于提供一种结合用户评价与标注的交互式图像检索方法，该方法通过用户对查询图像、想查找图像的文本描述或检索结果进行满意度评价，系统能够自动在线生成图像库的语义信息数据库，在检索的过程中，综合利用离线生成的图像底层物理特征和在线生成的高层语义信息进行检索，从而极大提高图像检索的准确度。同时，在生成语义信息数据库的过程中，充分考虑了不同用户给出的文本标注的可靠性因素。该方法具有效率高、精确度高、交互方式友好的特点。

[0006] 本发明的总体思想如下：用户提交一幅查询图像或对想查找图像进行文本描述或是二者相结合，检索系统会提取图像特征、提取文本描述中的关键字或者二者同时提取，进行检索时，系统会分别采用基于特征相似度匹配、基于文本相似度匹配和基于特征和文本相似度相结合匹配进行检索。然后用户可对检索结果进行“满意”或“不满意”的相关评价，通过用户对检索结果的相关评价，图像检索系统自动对满意图像进行关键字标注。该图像检索系统随着用户的不断使用，自动对图像库中的各图像生成丰富的语义信息数据库。本发明考虑到不同用户给出的文本标注可靠性不同的特点，在生成语义信息数据库的过程中，结合了用户的可信度。

[0007] 具体创新点：本发明利用基于物理特征和文本相结合的综合检索方法，从而极大提高图像的检索准确度。同时，在进行语义标注的过程中，只需对查询图像或想查找的图像

进行语义标注，图像检索系统自动对用户标记为满意的图像也进行语义标注，且产生的语义标注信息考虑到了用户的可信度因素。本发明提出的方法能通过本图像检索系统的不断使用，自动丰富图像库中各图像的语义标注信息，避免人工标注费时费力的问题。

[0008] 本发明提出的检索方法，具体步骤包括：(1) 对图像库中的所有图像进行特征提取，离线生成图像特征库 V；(2) 用户登录，如果不登录即为匿名用户；(3) 用户提交查询信息，该信息可以是查询图像  $Q_N$ ，也可以是输入的文本信息，或者是查询图像与文本信息同时存在；(4) 检索系统对查询图像  $Q_N$  进行特征提取，并对输入的文本信息提取关键字；(5) 检索系统对查询图像  $Q_N$  进行特征相似度匹配 S1；(6) 检索系统对提取的关键字进行文本的匹配，即语义相关性 S2；(7) 检索系统利用 S1 和 S2 计算总相似度 S，返回 S 值最大的图像集作为初始的检索结果集合 set(R)；(8) 同时，检索系统在关键字数据表 KL 中查询是否存在 set(R) 中各图像名对应的记录，如果存在，则给出这些记录对应的关键字权值排序最高的前 N( $N > 0$ ) 个关键字供用户选择；(9) 用户选择系统提供的关键字 K，同时也可以对图像文本标注  $T'$ ；(10) 系统根据用户选择的关键字和标注的文本信息，组合新关键字  $K'$ ；(11) 同时，用户对结果图像进行“满意”或“不满意”评价，对标记为“满意”的图像，用关键字  $K'$  进行文本标注；(12) 图像检索系统计算该用户的用户可信度 a，并计算“满意”图像标注的关键字权值 w，保存到关键字信息表 KL 中，同时，将注册用户的用户信息保存到用户信息数据表 UL 中，匿名用户的用户信息不需要保存，只需对 KL 进行更新；(13) 如果用户对检索结果不满意，则可修改文本标注 T，重复步骤 (3)–(12)，直到用户对检索结果满意为止，并给出最终的检索结果。

[0009] 更进一步，步骤 1、2 所述的特征提取中的特征包括：颜色特征、纹理特征、形状特征、空间分布特征及局部不变性特征。

[0010] 更进一步，步骤 6 所述的语义相关性 S2 的计算方法为：对于图像库中的每幅图像，在关键字信息数据表 KL 中查询包含该图像名及 K 中任一关键字对应的记录，从这些记录的关键字权值中选取最大值作为 S2，如果 KL 中没有该图像名与 K 中各关键字对应的记录，则 S2 值为 0。

[0011] 更进一步，步骤 7 所述的总相似度 S 的计算方法为： $S = \alpha S_1 + \beta S_2$ ，其中  $\alpha + \beta = 1$ ， $\alpha$  是基于特征相似度匹配在综合检索中所占的比重， $\beta$  是基于文本匹配在综合检索中所占的比重；当  $\alpha = 0$  时，综合检索实际是进行基于文本匹配的检索；当  $\beta = 0$  时，综合检索实际是进行基于特征相似度匹配的检索；当  $\alpha \neq 0$  且  $\beta \neq 0$  时，图像检索系统进行综合检索，根据用户的不同需求调整视觉相似度 S1 和语义相关性 S2 的比例；如果用户希望检索得到语义上更相关的图像，则调小  $\alpha$ ，调大  $\beta$ ，反之如果用户更需要视觉上相似的图像，则相应调小  $\beta$ ，调大  $\alpha$ 。

[0012] 更进一步，步骤 8、12 所述的关键字信息数据表 KL 的字段为：

[0013] 关键字、图像名、关键字评价次数、关键字权值

[0014] 其中 KL 中的一条记录仅与某一关键字及用该关键字表示的一幅图像的图像名一一对应，“关键字评价次数”为所有用户对该图像名用该关键字进行标注的次数的统计，“关键字权值”是每次不同用户用该关键字标注该图像时累加当前用户的可信度得到的。

[0015] 更进一步，步骤 12 中所述的用户信息数据表 UL 的字段为：

[0016] 用户名、密码、用户评价次数、用户可信度

[0017] 其中 UL 用来保存注册用户的操作记录，“用户评价次数”是统计该注册用户对各查询图像进行标注的次数总和，“用户可信度”用于表示该注册用户当前给出的文本标注的可信度。

[0018] 更进一步，步骤 12 中所述的 KL 和 UL 的更新，具体规则如下：

[0019] 1)KL 的更新：若某次检索过程取得的 K 中的某个关键字与用户标注为满意的某个图像名的对应记录不存在，则将该关键字、该图像的图像名、关键字评价次数（1 次）和该关键字对应该图像名的关键字权值作为一条记录存入 KL 中，若记录存在，则将该条记录的评价次数加 1，并重新计算该关键字对应记录的关键字权值项；

[0020] 2)UL 的更新：若进行检索的是注册用户，则将该用户对应记录的用户评价次数加 1，并重新计算该注册用户的用户可信度。

[0021] 更进一步，步骤 12 所述的用户可信度  $a$ ，其特征在于，对于匿名用户，它的可信度值只取决于本次操作；而对于注册用户，它的用户可信度是该用户多次评价的用户可信度的平均值；用户某次操作的用户可信度  $a$  的具体计算方法为：从 KL 中找到包含 K（用户选择的关键字以及用户输入的文本信息整合后的关键字集合）中任一关键字的所有记录，从这些记录中找到关键字权值最高的记录， $a$  的大小等于该条记录对应的关键字评价次数在对查询图像进行标注的所有关键字评价次数总和中所占比重对应的量化值；

[0022] 1) 对于注册用户，注册用户的第  $n$  次操作的用户可信度的平均值为：

$$[0023] a_n = \frac{a_{n-1} * (n - 1) + a}{n};$$

[0024] 2) 对于匿名用户，匿名用户的用户可信度值为  $a$ 。

[0025] 更进一步，在计算某次用户可信度时，关键字的评价次数在对  $Q_N$  进行标注的所有关键字评价次数总和中所占比重对应的量化值的量化方法可以是均匀量化，也可以是非均匀量化。

[0026] 更进一步，关键字信息数据表 KL 中的“关键字权值”字段，当前用户的可信度越高，则该关键字的准确性越高，该关键字对应的关键字权值也越大，某一关键字记录中的关键字权值计算方法如下：

[0027]

关键字权值 =  $\sum_{i=1}^m$  对该图像采用该关键字进行第  $i$  次评价的用户可信度， $m$  为该关键字的关键字评价次数。

## 附图说明

[0028] 图 1 图像检索系统框图

[0029] 图 2 图像检索在线检索流程图

## 具体实施方式

[0030] 下面结合附图对本发明作进一步说明：

[0031] 本实施例的技术方案如图 1 所示：

[0032] 首先,检索系统事先离线对图像库中的所有图像进行底层物理特征提取,提取的特征包括:颜色特征、纹理特征、形状特征、空间分布特征及局部不变性特征,生成图像特征库 V;然后进行如图 2 所示的在线检索过程,在线检索过程的具体流程如下:

[0033] (1) 用户提交查询信息,该信息可以是查询图像  $Q_N$ ,也可以是输入的文本信息,或者是查询图像与文本信息同时存在;检索系统对查询图像  $Q_N$  进行特征提取,并对输入的文本信息提取关键字;系统对查询图像进行特征相似度匹配 S1,对提取的关键字进行文本的匹配,即语义相关性 S2;检索系统利用 S1 和 S2 计算总相似度 S,  $S = \alpha S_1 + \beta S_2$ ;并对 S 值按从大到小进行排序,返回 S 值最大的图像集作为初始的检索结果集合 set(R);其中上述所说的公式  $S = \alpha S_1 + \beta S_2$ ,  $\alpha$ 、 $\beta$  分为三种情况:

[0034] (a) 用户只提交  $Q_N$ ,不进行任何文本标注,则相应的  $\beta = 0$ ,图像库中每幅图像的总相似度  $S = \alpha S_1$ ,综合检索实际是基于特征相似度匹配的检索;

[0035] (b) 用户既提交  $Q_N$ ,也对  $Q_N$  或想查找的图像进行相应的文本标注,则  $\alpha \neq 0$  且  $\beta \neq 0$ ,综合检索是基于特征相似度匹配和基于文本匹配相结合的综合检索;

[0036] (c) 用户不提交  $Q_N$ ,只对想查找的图像进行文本标注,则相应的  $\alpha = 0$ ,图像库中每幅图像的总相似度  $S = \beta S_2$ ,综合检索实际是基于文本匹配的检索。

[0037] (2) 图像检索系统在关键字数据表 KL 中查询是否存在 set(R) 中各图像名对应的记录,如果存在,则给出这些记录对应的关键字权值排序最高的前 N( $N > 0$ ) 个关键字供用户选择。

[0038] (3) 用户选择系统提供的关键字或者输入文本信息,系统从输入的文本信息中提取关键字并与用户选择的关键字合成新的关键字集合 K。

[0039] (4) 用户对 set(R) 上的图像进行“满意”或“不满意”评价,选择其认为比较接近检索目标的若干幅图像进行“满意”标记,形成相似图像集 S(M),系统自动把新关键字集合 K 标记相似图像集。

[0040] (5) 图像检索系统自动计算本次操作的用户可信度 a,并更新语义信息数据库中的用户信息数据表 UL:根据用户本次标注得到的关键字集合 K,从 KL 中找到包含 K 中任一关键字的所有记录,从这些记录中找到关键字权值最高的记录,得到该记录中对应的关键字评价次数,计算该关键字评价次数在对 set(R) 中的第一幅图像 R1 进行标注的所有关键字的关键字评价次数总和中所占比重,对该比重进行量化,将该量化值作为本次操作的用户可信度值 a。现给出 a 的计算方法,举例如下:假设 KL 中的存在 4 条对 R1 进行标注的关键字信息记录,对关键字评价次数所占比重的量化方法如下表所示:

[0041]

关键字	图像名	关键字评价次数	比重	排名	比重量化值	用户可信度
key1	R1	27	27/65	1	1	100%
key2	R1	12	12/65	3	0.6	60%
key3	R1	19	19/65	2	0.8	80%
key4	R1	5	5/65	4	0.4	40%

[0042] 假定本次查询,用户给出的关键字集合 K 为 key1 和 key3,系统找到对应的权值最高的记录,即关键字 key1 的记录,得到 key1 的关键字评价次数所占的比重为 27/65,根据该表给出的量化方法,得到比重量化值为 1,即得到本次操作的用户可信度 a 为 100%。对于匿名用户,匿名用户的此次操作的用户可信度大小就等于 a;对于注册用户,该注册用户此次操作的用户可信度为该用户进行过的评价次数 n 的平均值,即:

$$[0043] \quad a_n = \frac{a_{n-1} * (n - 1) + a}{n};$$

[0044] n 表示此次操作是该注册用户的第 n 次操作,图像检索系统将该注册用户的用户评价次数加 1,并更新用户可信度的值为  $a_n$ 。

[0045] (6) 根据本次操作的用户可信度,重新计算 K 中的每个关键字与相似图像集 S(M) 中每幅图像对应的关键字权值 w, w 的更新方法为:

$$W = \sum_{i=1}^m \text{对该图像采用该关键字进行第 } i \text{ 次评价的用户可信度}, \text{ 即累加上本次操作的}$$

用户可信度;图像检索系统自动查询 S(M) 中各图像的图像名与 K 中各关键字在 KL 中是否存在对应的记录,如果存在,则将该图像名与该关键字对应的关键字评价次数加 1,对应的关键字权值加上本次操作的用户可信度 a;如果不存在,则将该图像的图像名、该关键字、关键字评价次数(1 次)和相应关键字权值(关键字权值即为本次操作的用户可信度值 a)作为一条记录添加进 KL 中。

[0046] (7) 如果用户对检索结果不满意,则用户可修改文本标注 T,并重新进行基于特征相似度匹配和基于文本匹配相结合的综合检索,直到得到用户满意的结果为止。

[0047] 应当理解的是,上述针对实施实例的描述较为具体,并不能因此而认为是对本发明专利保护范围的限制,本发明的专利保护范围应以所附权利要求为准。

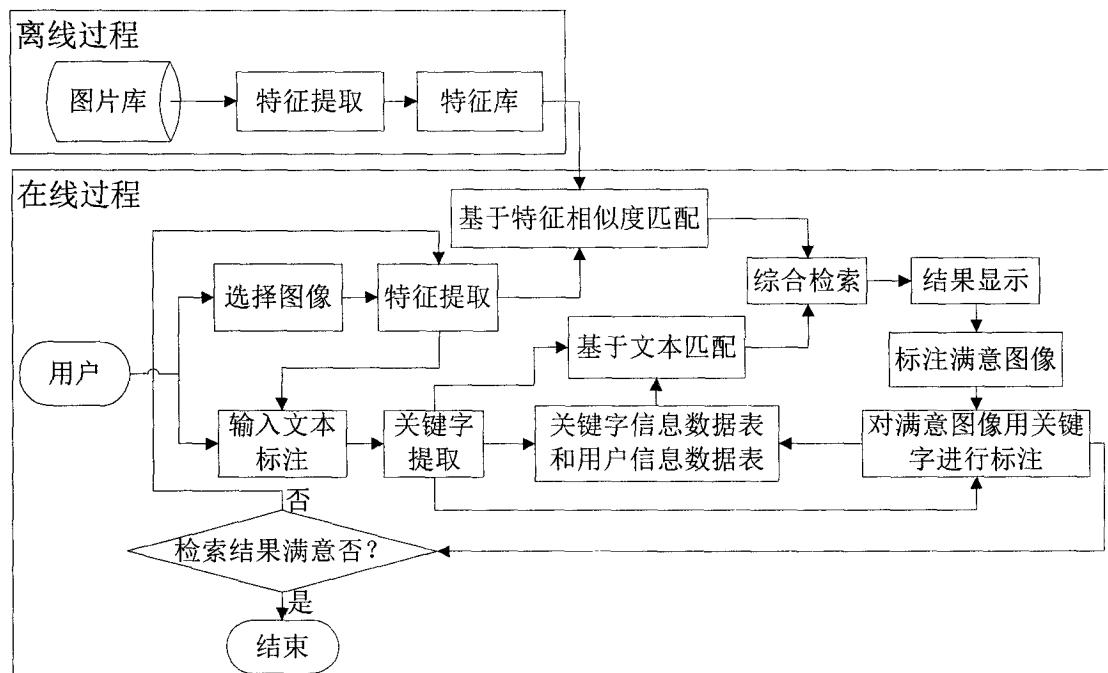


图 1

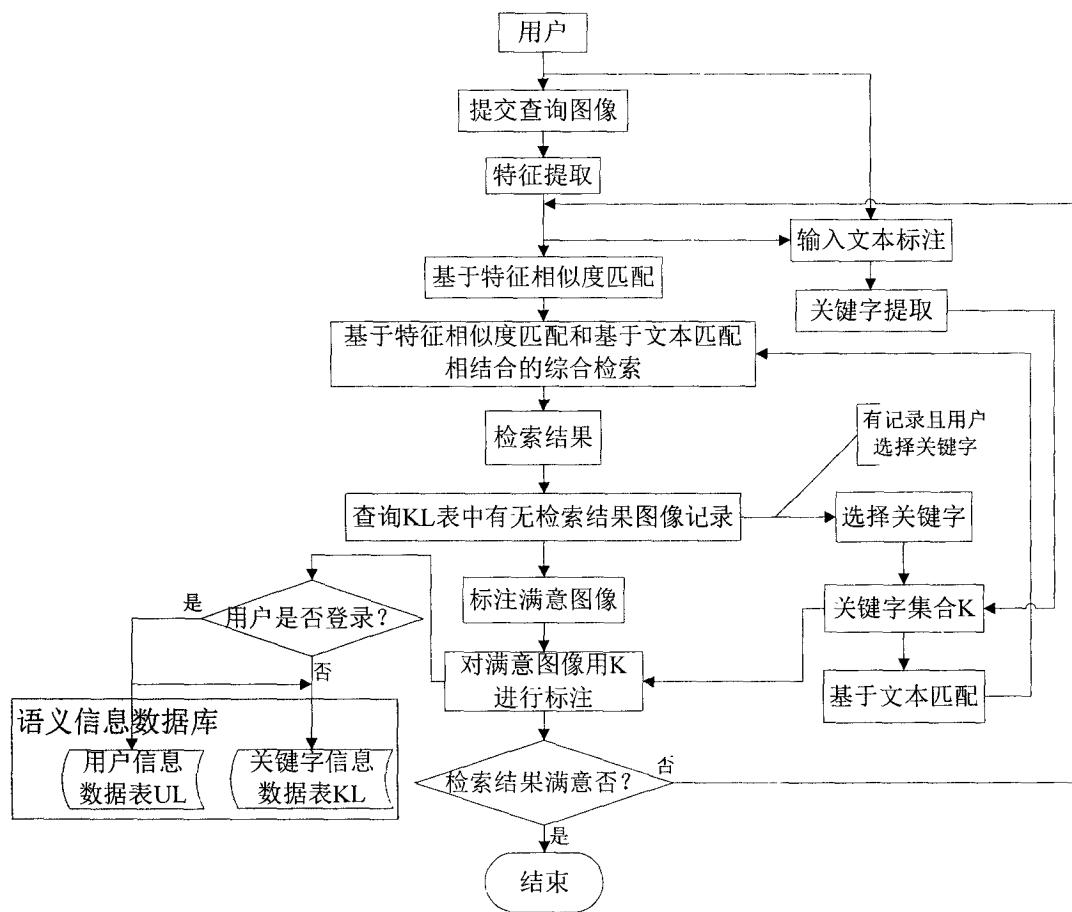


图 2