

(12) 发明专利

(10) 授权公告号 CN 102521227 B

(45) 授权公告日 2013. 07. 31

(21) 申请号 201110317167. 7

(22) 申请日 2011. 10. 18

(73) 专利权人 中国科学院自动化研究所  
地址 100190 北京市海淀区中关村东路 95 号

(72) 发明人 徐常胜 桑基韬

(74) 专利代理机构 中科专利商标代理有限责任  
公司 11021

代理人 梁爱荣

(51) Int. Cl.  
G06F 17/30 (2006. 01)

审查员 马晓宇

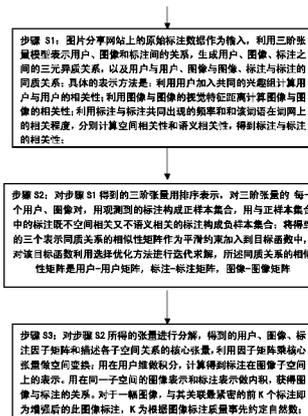
权利要求书1页 说明书6页 附图3页

(54) 发明名称

基于用户信息建模的图像标注增强方法

(57) 摘要

本发明是一种基于用户信息建模的图像标注增强方法,首先,用户、图像及标注的原始关系被表示成一个三阶张量;然后,观测到的原始标注被表示成正负样本对,张量分解的目标是最小化正负样本对的排序关系;最终,利用分解得到的核心张量和因子矩阵进行子空间变换,可以得到共同子空间上的图像及标注表示。在大规模图片分享网站数据上,证明了发明的方法获得了更好的标注增强效果。



1. 一种基于用户信息建模的图像标注增强方法,其特征在于,该方法的步骤包括:

步骤 S1:图片分享网站上的原始标注数据作为输入,利用三阶张量表示用户、图像和标注间的关系,生成用户、图像、标注之间的三元异质关系,以及用户与用户、图像与图像、标注与标注的同质关系;具体的表示方法是:利用用户加入共同的兴趣组计算用户与用户的相关性;利用图像与图像的视觉特征距离计算图像与图像的相关性;利用标注与标注共同出现的频率和该标注词语在词网 WordNet 上的相关程度,分别计算标注与标注的空间相关性和语义相关性;

步骤 S2:对三阶张量用排序表示,对三阶张量的每一个用户对、图像对,用观测到的原始标注构成正样本集合,用与正样本集合既不空间相关又不语义相关的标注构成负样本集合;将同质关系的相似性矩阵作为平滑约束加入到目标函数中,利用选择优化方法对该目标函数进行迭代求解,得到的用户、图像、标注因子矩阵和描述各子空间关系的核心张量;

步骤 S3:利用因子矩阵乘核心张量,对用户、图像、标注因子矩阵和描述各子空间关系的核心张量做空间变换;在用户维做积分,计算得到标注在图像子空间上的表示;用在同一子空间的图像表示和标注表示做内积,获得图像与标注的关系;对于一幅图像,与该幅图像关联最紧密的前 K 个标注即为增强后的该幅图像标注, K 为根据该幅图像标注质量事先约定自然数。

2. 如权利要求 1 所述的基于用户信息建模的图像标注增强方法,其特征在于,利用三阶张量对标注信息的用户、图像、标注关系进行计算和建模,将用户信息加入到图像与标注的分析和增强任务中。

3. 如权利要求 1 所述的基于用户信息建模的图像标注增强方法,其特征在于,图像标注的增强任务被转化为根据原始标注张量进行重构,最小化目标函数的过程。

4. 如权利要求 3 所述的基于用户信息建模的图像标注增强方法,其特征在于,目标函数是由对原始标注生成的正负样本对排序关系的破坏以及由同质关系相似度矩阵构建的正则化项组成。

5. 如权利要求 1 所述的基于用户信息建模的图像标注增强方法,其特征在于,通过计算同一子空间内图像和标注向量的内积获得图像增强后的标注;标注矩阵需要先乘核心矩阵再对用户维积分,以转换为图像子空间的向量计算。

## 基于用户信息建模的图像标注增强方法

### 技术领域

[0001] 本发明属于信息检索技术领域,涉及基于内容的图像检索,具体地涉及用户建模及其在图像标注增强的方法。

### 背景技术

[0002] 基于内容的图像检索(CBIR)经历了以图找图(query by example)、基于图像周围文本信息(query by surrounding text)和基于标注信息(query by tag)三个阶段,其中基于标注信息的图像检索是真正的基于语义的,被认为最有希望解决语义鸿沟的问题。随着媒体分享网站的兴起,用户在上传图片的过程中,会附带为图片提供标注;但这些原始标注信息存在着不准确、不完整和主观模糊等问题,影响了图像检索的效率。图像标注增强旨在对原始标注进行处理,去除噪音、补充缺失标注,使其更好地为图像检索任务服务。

[0003] 图像标注增强方法实际就是分析图像和标注之间的关系,目前的标注增强方法利用第  $n$  图像与第  $n+1$  图像、第  $n$  标注与第  $n+1$  标注的相似度分析传播影响重构图像和标注之间的关系,采用的方法包括随机行走、稀疏分解、概率矩阵分解等等。然而,原始标注是由用户产生的,原始标注的不准确、不完整和主观模糊问题也是用户的行为模式带来的,现有方法忽略了用户信息在标注过程中的作用。

### 发明内容

[0004] (一)要解决的技术问题

[0005] 为了解决现有技术方案存在技术问题,本发明的目的是提供一种基于用户信息建模的图像标注增强方法。

[0006] (二)技术方案

[0007] 为实现上述目的,本发明提供一种基于用户信息建模的图像标注增强方法,该方法的步骤包括:

[0008] 步骤 S1:图片分享网站上的原始标注数据作为输入,利用三阶张量表示用户、图像和标注间的关系,生成用户、图像、标注之间的三元异质关系,以及用户与用户、图像与图像、标注与标注的同质关系;具体的表示方法是:利用用户加入共同的兴趣组计算用户与用户的相关性;利用图像与图像的视觉特征距离计算图像与图像的相关性;利用标注与标注共同出现的频率和该词语在词网上的相关程度,分别计算标注与标注的空间相关性和语义相关性;

[0009] 步骤 S2:对三阶张量用排序表示,对三阶张量的每一个用户对、图像对,用观测到的原始标注构成正样本集合,用与正样本集合既不空间相关又不语义相关的标注构成负样本集合;将同质关系的相似性矩阵作为平滑约束加入到目标函数中,利用选择优化方法对该目标函数进行迭代求解,得到的用户、图像、标注因子矩阵和描述各子空间关系的核心张量;

[0010] 步骤 S3:利用因子矩阵乘核心张量,对用户、图像、标注因子矩阵和描述各子空间

关系的核心张量做空间变换；在用户维做积分，计算得到标注在图像子空间上的表示；用在同一子空间的图像表示和标注表示做内积，获得图像与标注的关系；对于一幅图像，与该幅图像关联最紧密的前  $K$  个标注即为增强后的该幅图像标注， $K$  为根据该幅图像标注质量事先约定自然数。

[0011] 优选实施例，利用三阶张量对标注信息的用户、图像、标注关系进行计算和建模，将用户信息加入到图像与标注的分析和增强任务中。

[0012] 优选实施例，图像标注的增强任务被转化为根据原始标注张量进行重构，最小化目标函数的过程。

[0013] 优选实施例，目标函数是由对原始标注生成的正负样本对排序关系的破坏以及由同质关系相似度矩阵构建的正则化项组成。

[0014] 优选实施例，通过计算同一子空间内图像和标注向量的内积获得图像增强后的标注；标注矩阵需要先乘核心矩阵再对用户维积分，以转换为图像子空间的向量计算。

[0015] （三）本发明的有益效果：

[0016] 针对原始标注进行标注增强就成为了基于标注的图像检索和理解的重要步骤。本发明在图像标注分析中引入用户信息，提出了基于排序的多关联张量分解的方法，可以得到更紧致的图像和标注表示，从而进行图像标注的增强。本发明解决了用户产生的大量主观模糊、不完整不准确的原始标注的问题。用户信息的引入可以帮助理解原始标注中的主观模糊、不准确等问题，并进一步分解原始的图像和标注矩阵，得到更紧致的图像和标注表示，从而用于分析图像和标注之间的关系。

## 附图说明

[0017] 图 1a 是本发明用户信息的图像标注增强方法的流程图；

[0018] 图 1b 是本发明用户信息的图像标注增强框架图；

[0019] 图 2a- 图 2b 是本发明的标注数据的表示示意图，其图 2a 显示的是图 1b 中的例子构建的；

[0020] 图 3 是本发明的张量 (Tucker) 分解示意图；

[0021] 图 4a- 图 4f 是本发明实施例图像增强结果示例。

## 具体实施方式

[0022] 为使本发明的目的、技术方案和优点更加清楚明白，以下结合具体实施例，并参照附图，对本发明进一步详细说明。

[0023] 本发明考虑到用户信息在图像标注过程中的作用，本发明实现了一种基于用户信息建模的图像标注增强方法。本发明的方法充分利用了图像分享网站中用户、图像、标注三者之间的各种交互行为，从因子分析的角度重构用户、图像及标注之间的关系，并得到紧致的图像和标注表示，从而用于图像的标注增强。在大规模网络数据集上，通过与传统的图像标注增强方法比较，本发明的方法可以获得更好的标注增强效果。

[0024] 本发明考虑用户信息的图像标注增强的结构流程图在图 1a 和图 1b 中显示，其包含三个组成部分：1) 数据采集和表示，例如图 1b 中，图像的原始标注中包括飞机、天空、II、外星人、棒极了等，表示用户对该幅图像的理解；2) 对基于排序的多关联张量的张量分解

方法 ;3) 图像标注增强 ;图 1b 中包括多个用户,多个标注以及多个图像,u 代表用户,t 代表标注,i 代表图像,描述 u、t、i 的个数时用下标  $n = 1, 2, 3, 4, \dots$  分别代表用户  $u_n$ 、标注 t、图像  $i_n$  的序数,例如  $u_1$  表示第一个用户, $i_2$  表示第二个图像, $t_3$  表示第三个标注,等等。

[0025] 1. 数据采集和表示

[0026] 多个用户的多个图像和多个标注在图片分享网站上存在着丰富的交互信息,本发明要利用的主要有两类信息 :1) 一个用户的一个图像及一个标注之间的异质关系,即用户对图像进行了哪些标注 ;2) 第一用户  $u_1$  与第二用户  $u_2$ 、第一图像  $i_1$  与第二图像  $i_2$ 、第一标注  $t_1$  与第二标注  $t_2$  的同质关系,比如多个用户共同加入了哪些兴趣组,多个图像之间的视觉相似度,多个标注之间的语义相关性等。

[0027] 1.1 用户、图像、标注的异质关系

[0028] 标注行为包含了用户、图像和标注三类实体,所述用户实体、图像实体和标注实体分别用集合符号  $\Omega, \Lambda, \Gamma$  表示。用户的图像和标注的异质关系描述的实际就是用户的原始标注,观测到的异质关系可表示为  $O \subset \Omega \times \Lambda \times \Gamma$ ,即任何一个三元组  $(u, i, t) \in O$  意味着用户 u 对图像 i 标注 t 了。因此,这种三阶关系可以很自然地用三阶张量来表示,用户 u、图像 i、标注 t 的三阶关系为  $Y \in R^{|\Omega| \times |\Lambda| \times |\Gamma|}$  可表示如下 :

[0029]

$$y_{u,i,t} = \begin{cases} 1 & \text{若 } (u, i, t) \\ 0 & \text{其他} \end{cases} \quad (1)$$

[0030] 其中  $|\Omega|, |\Lambda|, |\Gamma|$  分别是用户实体、图像实体和标注实体的数量,R 表示实数集合, $y_{u,i,t}$  表示用户 u、图像 i、标注 t 的三阶关系。图 2a- 图 2b 是本发明的标注数据的表示示意图,其图 2a 显示的是图 1b 中的例子构建的张量。我们假设系统里共有三个用户、四幅图像和五个可供选择标注,最外面的  $4 \times 5$  矩阵表示的就是第一用户  $n_1$  的标注情况,如第一列表示第一用户  $n_1$  对第一图像  $i_1$  加了第三标注  $t_3$ 。

[0031] 1.2 用户、图像、标注的同质关系

[0032] 除了用户、图像及标注之间的异质关系,本发明还利用了第一用户  $u_1$  与第二用户  $u_2$ 、第一图像  $i_1$  与第二图像  $i_2$ 、第一标注  $t_1$  与第二标注  $t_2$  的同质关系,分别构建了用户、图像和标注的相似矩阵 (affinity matrix) 为  $W^U \in R^{|\Omega| \times |\Omega|}$ ,  $W^I \in R^{|\Lambda| \times |\Lambda|}$ ,  $W^T \in R^{|\Gamma| \times |\Gamma|}$ 。

[0033] 用户加入共同的兴趣组可以反映用户的背景以及兴趣信息,因此本发明利用这一信息计算用户 - 用户的同质关系  $W_{m,n}^U$  :

$$W_{m,n}^U = \frac{N(u_m, u_n)}{N(u_m) + N(u_n)} \quad (2)$$

[0035] 其中  $N(u_m)$ ,  $N(u_n)$  分别是用户  $u_m$ ,  $u_n$  加入的兴趣组数目,  $N(u_m, u_n)$  是用户  $u_m$  和  $u_n$  共同加入的兴趣组数目,  $m, n$  是小于  $|\Omega|$  的自然数。

[0036] 为计算第 n 图像与第 n+1 图像的视觉相似度,本发明每幅图像抽取了 428 维的特征向量 d,包括 225 维分块颜色特征、128 维小波纹理特征以及 75 维边缘分布直方图特征。图像相似矩阵  $W_{m,n}^I$  按如下公式计算 :

$$W_{m,n}^I = e^{-\|d_m - d_n\|^2 / \sigma_I^2} \quad (3)$$

[0038] 其中第  $m$ 、 $n$  个 428 维的特征向量表示为  $d_m$ 、 $d_n$ ， $\sigma_1$  是初始矩阵  $W^1$  的中位数。

[0039] 为同时考虑标注之间的语义相关性 (semantic relevance) 和空间相关性 (context relevance)，标注的相似度矩阵计算分为如下两部分。与计算用户相似矩阵类似，标注  $t_n$  与标注  $t_m$  之间的空间相关性由它们的共同出现频率计算：

$$[0040] \quad t_{m,n}^c = \frac{n(t_m, t_n)}{n(t_m) + n(t_n)} \quad (4)$$

[0041] 标注的语义相关性通过词网 (WordNet) 距离计算：

$$[0042] \quad t_{m,n}^s = \frac{2 \cdot IC(lcs(t_m, t_n))}{IC(t_m) + IC(t_n)} \quad (5)$$

[0043] 其中  $t_{m,n}^c$ 、 $t_{m,n}^s$  分别代表空间相关性和语义相关性， $IC(\cdot)$  是标注的信息熵， $lcs(t_m, t_n)$  是标注  $t_m$  和标注  $t_n$  在词网 (WordNet) 上的最近公共父词，上标  $c$ 、 $s$  分别代表空间相关性和语义相关性。最终的标注相似矩阵  $W_{m,n}^T$  按如下计算：

$$[0044] \quad W_{m,n}^T = \lambda_c t_{m,n}^c + \lambda_s t_{m,n}^s \quad (6)$$

[0045] 其中  $\lambda_c$ 、 $\lambda_s$  是权重系数，且  $\lambda_c + \lambda_s = 1$ 。

[0046] 2 基于排序的多关联张量分解方法

[0047] 2.1 因子矩阵表示和张量分解

[0048] 由原始标注构建的张量  $Y$ ，本发明的任务是要在重构后的张量的过程中得到标注和图像的表达。从因子分析的角度，张量分解得到的标注和图像因子矩阵就是所需要的表示。本发明采用的张量分解方法是塔克 (Tucker) 分解：

[0049]

$$\hat{Y} := C \times_u U \times_i I \times_t T \quad (7)$$

[0050] 其中  $\hat{Y}$  表示重构后的张量， $\times_u$ 、 $\times_i$ 、 $\times_t$  分别表示张量在用户、图像、标注维度乘矩阵。每一个低秩矩阵  $U$ 、 $I$ 、 $T$  分别对应于一个因子表示，核心张量  $C \in R^{r_u \times r_i \times r_t}$  包含着不同因子之间的联系， $r_u$ 、 $r_i$ 、 $r_t$  分别表示分解后核心张量的用户、图像和标注维度，如图 3。

[0051] 2.2 基于排序的优化框架

[0052] 传统的直接根据公式 (1) 进行张量低秩近似的分解方法不适于图像标注增强问题。一方面，用户不可能为图像添加所有相关的标注；另一方面，由于标注的不完整性，未观测到的标注统一用 0 填充是不合理的。因此，本发明提出了一种基于排序的优化框架进行张量分解。首先，优化方程中没有必要回归到绝对的 0 或 1，相比于分类，标注增强问题关注的排序，即标注的顺序关系。如  $y_{u,i,t^+} > y_{u,i,t^-}$  的意思是用户  $u$  认为标注  $t^+$  相比标注  $t^-$  更适合用来描述图像  $i$ ，其中  $y_{u,i,t^+}$ 、 $y_{u,i,t^-}$  表示三阶张量的两个元素。对于每一个用户 - 图像对  $(u, i)$ ，借助排序学习 (learning-to-rank) 的思想，本发明的任务就是寻找这样的样本对 (pair-wise) 进行训练。

[0053] 本发明定义一个用户  $u$  和图像  $i$  的二元组  $(u, i)$  为一个位 (post) 观测到的位组成集合  $P_0$ ：

$$[0054] \quad P_0 = \{(u, i) \mid \exists t \in \Gamma, y_{u,i,t} = 1\} \quad (8)$$

[0055] 其中  $y_{u,i,t}$  表示张量  $Y$  在  $(u, i, t)$  位置的元素,中性的三元组组成集合  $M$  :

$$[0056] \quad M = \{(u, i, t) \mid (u, i) \notin P_o\} \quad (9)$$

[0057] 中性的三元组既不适宜当做正样本也不适于用作负样本,因此被舍弃,在图 2b 中用问号填充。图 2b 是对图 2a 进行分析转换后的结果,其中问号表示舍弃不放入目标函数中的数据,加号表示训练数据的正样本,减号表示负样本。

[0058] 对于位  $(u, i) \in P_o$ ,观测到的标注组成正样本集  $\Gamma_{u,i}^+$  :

[0059]

$$\Gamma_{u,i}^+ = \{t \mid (u, i) \in P_o \wedge y_{u,i,t} = 1\} \quad (10)$$

[0060] 在选择负样本时,本发明考虑了标注的语义相关性和空间相关性,即与正样本集中标注既不语义相关又不空间相关的标注构成负样本集  $\Gamma_{u,i}^-$  :

[0061]

$$\Gamma_{u,i}^- = \{t \mid (u, i) \in P_o \wedge y_{u,i,t} \neq 1 \wedge t \notin N_{\Gamma_{u,i}^+}\} \quad (11)$$

[0062] 其中  $N_{\Gamma_{u,i}^+}$  表示与位  $(u, i)$  中的观测标注语义且空间不相关的标注集合。

[0063] 如图 2a 中,假设标注 1、标注 2 与标注 3 相关,则  $t_4, t_5 \in \Gamma_{u,i}^-$ ,如图 2b,正样本集的标注由加号表示,负样本集合的标注则由减号表示。

[0064] 假设对于任何标注  $t \in \Gamma_{u,i}^+$  比所有的标注  $t \in \Gamma_{u,i}^-$  都更适合于描述图像,即 :

[0065]

$$\hat{y}_{u,i,t_1} > \hat{y}_{u,i,t_2} \Leftrightarrow t_1 \in \Gamma_{u,i}^+ \wedge t_2 \in \Gamma_{u,i}^- \quad (12)$$

[0066] 其中  $\hat{y}_{u,i,t_1}$ 、 $\hat{y}_{u,i,t_2}$  表示重构后张量  $\hat{Y}$  在位置  $(u, i, t_1)$  和  $(u, i, t_2)$  的元素。优化目标就是在重构的张量  $\hat{Y}$  中尽可能地减少对训练样本对排序关系的破坏,可以表示如下 :

[0067]

$$\min_{U,I,T,C} \sum_{(u,i) \in P_o} \left( \sum_{t^+ \in \Gamma_{u,i}^+} \sum_{t^- \in \Gamma_{u,i}^-} f(\hat{y}_{u,i,t^-} - \hat{y}_{u,i,t^+}) \right) \quad (13)$$

[0068] 其中  $f$  是单调递增函数,值域为  $[0, 1]$ ,  $\min_{U,I,T,C}$  表示优化变量  $U, I, T, C$  以最小化目标函数,其中  $t^+$ 、 $t^-$  分别表示正、负标注集合中的标注,  $\hat{y}_{u,i,t^-}$ 、 $\hat{y}_{u,i,t^+}$  分别表示重构后张量  $\hat{Y}$  在位置  $(u, i, t^-)$  和  $(u, i, t^+)$  的元素。

[0069] 2.3 多关联平滑约束

[0070] 为解决训练样本的稀疏性问题,用户、图像、标注的相似度矩阵被作为平滑约束加入到公式 (13) 的优化方程中;再加入惩罚大参数  $\lambda$  正则化项,最终的优化方程  $\min_{U,I,T,C} g$  可表示为 :

$$[0071] \quad \min_{U,I,T,C} g = \sum_{(u,i) \in P_o} \left( \sum_{t^- \in \Gamma_{u,i}^-} \sum_{t^+ \in \Gamma_{u,i}^+} f(\hat{y}_{u,i,t^-} - \hat{y}_{u,i,t^+}) \right) + (14)$$

$$[0072] \quad \alpha \left( \text{tr}(U^T L_U U) + \text{tr}(I^T L_I I) + \text{tr}(T^T L_T T) \right) + \beta \left( \|U\|^2 + \|I\|^2 + \|T\|^2 \right)$$

[0073] 其中  $g$  是目标函数,  $U^T, I^T, T^T$  中的上标  $T$  表示矩阵的转置,  $\|\cdot\|^2$  表示矩阵的二范数,  $\text{tr}(\cdot)$  表示矩阵的秩,  $L_U, L_I, L_T$  分别表示矩阵  $U, I, T$  的拉普拉斯矩阵,  $\alpha, \beta$  是系数,控制正则化项的权重。

[0074] 公式 (14) 无法直接求解,可采用选择优化 (alternative optimization) 的方式

迭代求解。

### [0075] 3 图像标注增强

[0076] 通过求解公式 (14), 可以得到因子矩阵  $U, I, T$  以及核心张量  $C$ 。核心张量实际定义的是一种多线性变换, 描述子空间之间的关系, 用核心张量乘因子矩阵相当于做了空间变换。定义张量  $T^{UI} = C \times_t T$ , 则  $T^{UI} = R^{r_U \times r_I \times |I|}$  可认为是标注在图像  $X$  用户空间的表示, 每个  $r_U \times r_I$  矩阵表示一个标注。在用户维对  $T^{UI}$  积分, 就得到标注在图像空间的表示。因此, 跨空间的图像 - 标注关系矩阵  $X^{II}$  可由下式计算:

$$[0077] \quad X^{II} = I \bullet \left( C \times_t T \times_u 1_{r_U}^T \right) \quad (15)$$

[0078] 其中  $I$  是张量分解得到的图像因子矩阵,  $1_{r_U}$  为长度为  $r_U$  的单位向量, 对第  $m$  个图像, 前  $K$  个关系值最大的标注即为其增强后的标注  $Top(m, K)$ :

$$[0079] \quad Top(m, K) = \max_{t \in T}^K X_m^{II} \quad (16)$$

[0080] 其中  $X_m^{II}$  表示矩阵  $X^{II}$  的第  $m$  行元素。

### [0081] 实施效果

[0082] 为了评估本本发明, 本发明在真实的图像分享网站 Flickr 的数据集 NUS\_WIDE 上进行了实验。数据集包括了 27 万张图片, 这些图片来自 5 万个用户。数据集提供了 81 个标注的真实值, 本发明在这 81 个标注上评估了 F1 score, 结果比现有最好的图像增强算法提高了约 5%。

[0083] 图 4a- 图 4f 显示了利用本发明方法进行标注增强的结果示例。在每幅图像的左侧, 上面是原始标注, 下面是经过增强后的标注。对于图像图 4c 和图 4f, 图像本身很复杂, 依靠分析图像的视觉相关性来还原图像和标注的关系非常困难; 通过引入用户信息, 本发明发现标注图 4c 的用户同时对其他图像大量标注了类似 ' 拼图 (mosaic) ', ' 建筑 (building) ' 这类的词, 而图 4f 的标注者则是一名雕塑爱好者, 因而在增强的标注中, 这些新的语义被添加进来。又如图像图 4a, 在原始标注中, 只有 ' 机场 (airport) ' 与图像内容相关, 经过标注增强后, 主观的标注被过滤, 并引入了空间相关的标注 ' 大飞机 (airplane) ', ' 路 (road) ' 以及语义相关的标注 ' 飞机 (plane) ' ; 通过图像图 4e 以及图 4d 的结果也可以看出基于排序的优化框架的优势。图像图 4b 则表明了本发明提出的方法也适用于原始没有标注的图像, 即可以用于图像自动标注问题。

[0084] 以上所述, 仅为本发明中的具体实施方式, 但本发明的保护范围并不局限于此, 任何熟悉该技术的人在本发明所揭露的技术范围内, 可理解想到的变换或替换, 都应涵盖在本发明的包含范围之内, 因此, 本发明的保护范围应该以权利要求书的保护范围为准。

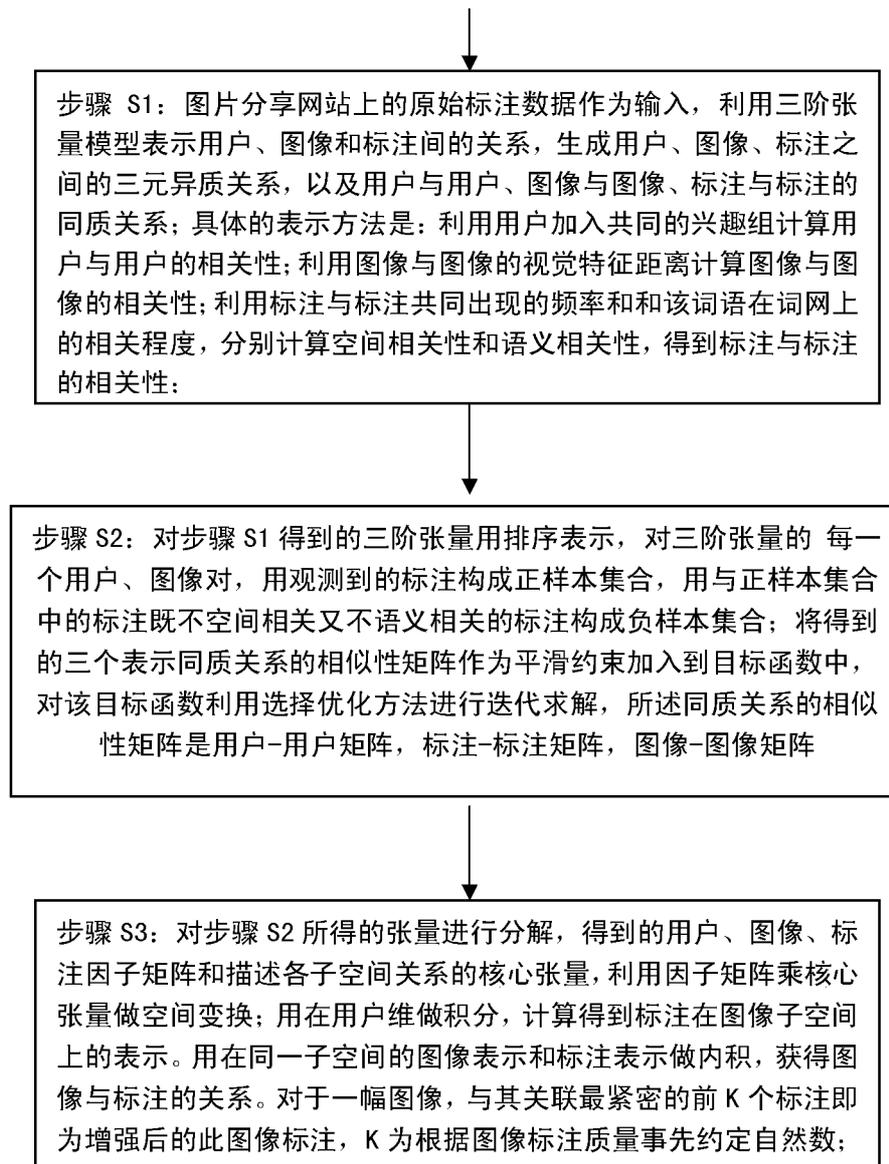


图 1a

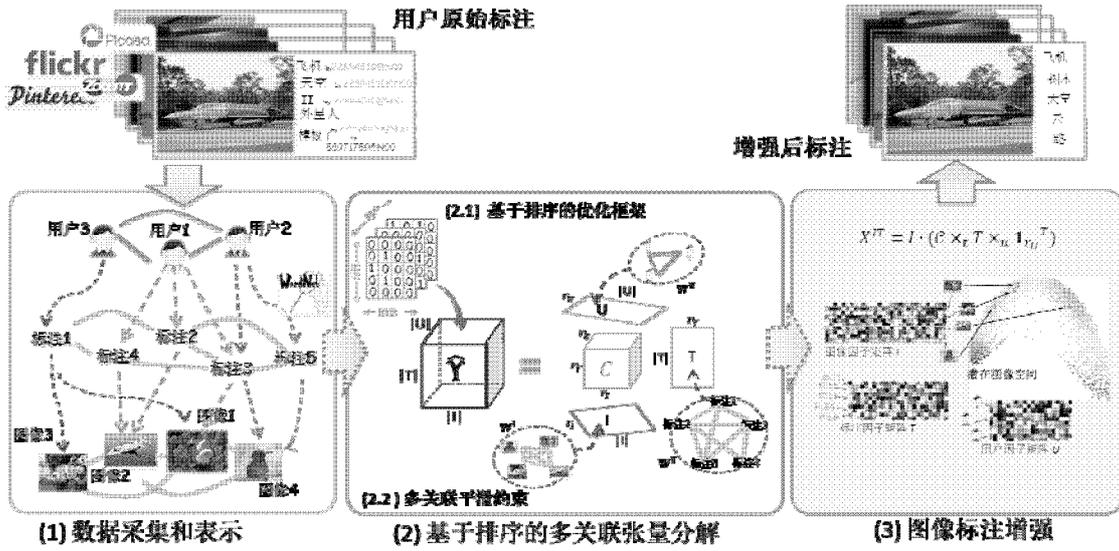


图 1b

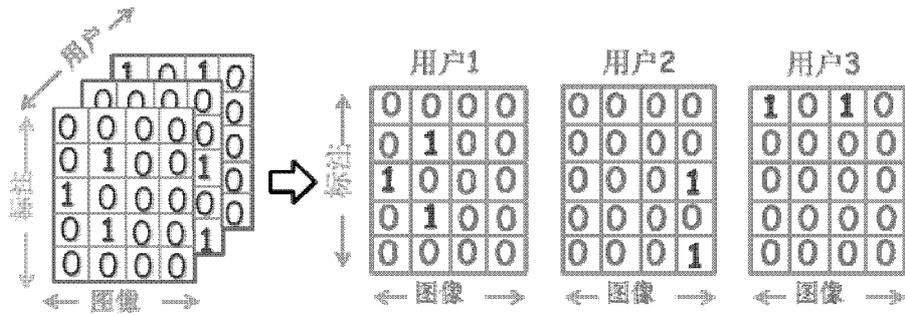


图 2a

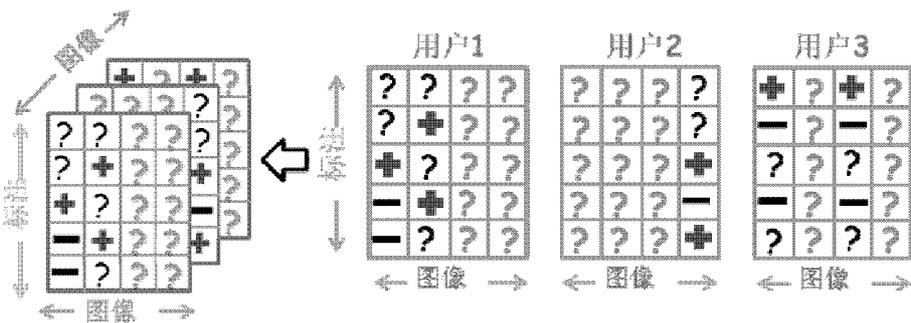


图 2b

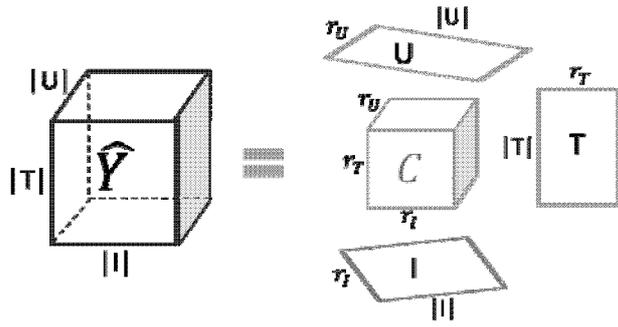
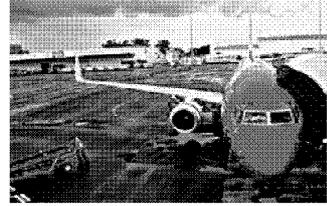


图 3

原始标注  
机场 737  
维珍蓝



增强后标注  
飞机 路 艺术  
机场 大飞机

图 4a

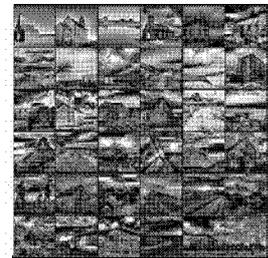
原始标注



增强后标注  
险 黑暗 猎豹  
猫 宠物

图 4b

原始标注



户外 2006  
遗弃物

增强后标注  
建筑 自然 户外  
黑暗 拼图

图 4c

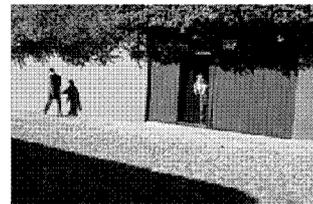
原始标注  
宠物 雌性 猫  
对话



增强后标注  
猫 白色 鱼  
水 宏观

图 4d

原始标注  
舌头 散步  
24小时



增强后标注  
散步 树 路  
行人 河

图 4e

原始标注  
加利福尼亚  
视觉



增强后标注  
建筑 水  
加利福尼亚 光线  
棒极了

图 4f