



(12) 发明专利申请

(10) 申请公布号 CN 103886066 A

(43) 申请公布日 2014.06.25

(21) 申请号 201410105511.X

(22) 申请日 2014.03.20

(71) 申请人 杭州禧颂科技有限公司

地址 311100 浙江省杭州市下城区费家塘路
588号4幢424室

(72) 发明人 陈晋音 黄坚

(74) 专利代理机构 杭州求是专利事务所有限公
司 33200

代理人 张法高

(51) Int. Cl.

G06F 17/30(2006.01)

G06K 9/66(2006.01)

权利要求书2页 说明书6页 附图2页

(54) 发明名称

一种基于鲁棒非负矩阵分解的图像检索方法

(57) 摘要

本发明公开了一种基于鲁棒非负矩阵分解的图像检索方法。该方法针对非负矩阵分解算法模型在图像间关系表达和抗噪两方面的不足,首先依据图像数据的视觉特征集 X 构建表征数据分布关系的 $L1$ 图约束,并加入稀疏噪声项,来构建鲁棒非负矩阵分解算法模型;接着在模型训练阶段,采用迭代优化求解方法,从 X 中得到图像特征基矩阵 U 和所有图像在 U 所展开的子空间中的特征表达 V ;最后在图像检索时,将检索参考的图像 x 转换到 U 的子空间,得到新特征表达 v ;依据 v 与图像数据集中图像的距离,对图像排序,返回距离最小的前 K 个图像给用户,实现对图像的检索功能。本方法具有较强的抗噪能力,对新图像的处理计算时间为线性时间,能够快速准确地进行图像检索。

1. 一种基于鲁棒非负矩阵分解的图像检索方法,其特征在于包括如下步骤:

1) 从公共图像数据集 Core15K 中选择 50 类共 5000 张图像,提取 PHOG 视觉特征集 $X = [x_1, \dots, x_n]$,其中 $n = 5000$, x_i 为第 i 张图像所对应的 PHOG 视觉特征向量;将 PHOG 视觉特征集 X 作为所要构建的 L1 图的顶点,L1 图的边权重矩阵 $W = [w_1, \dots, w_n]$,其中 w_i 计算如下:

$$\begin{aligned} \min \|w_i\|_1 \\ \text{s.t. } \|x_i - Xw_i\| \leq \varepsilon, w_{ii} = 0, \forall i \end{aligned} \quad 1$$

其中, x_i 为第 i 幅图像的视觉特征, w_i 为矩阵 W 的第 i 列, ε 为误差系数, w_{ii} 为向量 w_i 的第 i 个分量;

构建表征图像数据分布结构关系的 L1 图约束如下:

$$\begin{aligned} \min R_g = \sum_{i=1}^n \|v_i - Vw_i\|_F^2 = \|V - VW\|_F^2 \\ = \text{tr}(V(I-W)(I-W)^T V^T) = \text{tr}(VGV^T) \end{aligned} \quad 2$$

其中, $G = (I-W)(I-W)^T$, $V = [v_1, \dots, v_n]$ 为图像集在子空间上的特征表达, I 为单位矩阵;

2) 将 L1 图约束和稀疏噪声约束项 E 加入到传统的非负矩阵分解算法框架中,构建鲁棒非负矩阵分解算法模型:

$$\min_{U, V \geq 0, E} \|X - UV - E\|_F^2 + \lambda_1 \cdot \|E\| + \lambda_2 \cdot \|U\|_F^2 + \lambda_3 \cdot \text{tr}(VGV^T) \quad 3$$

其中, λ_1 , λ_2 和 λ_3 为正则化参数,在实验中分别设置为 0.01, 0.001 和 0.1, $U \in R^{d \times r}$ 为保持视觉和语义相似性的图像基矩阵, $V \in R^{r \times n}$ 为图像基矩阵 U 所展开的子空间中的特征表达, $E \in R^{d \times n}$ 是稀疏噪声约束项, d 为图像特征的维度, r 为低维特征空间的维度;

3) 利用图像数据集对公式 3 算法模型进行训练,采用迭代优化求解,计算得到保持视觉和语义相似性的图像基矩阵 U 和所有图像在图像基矩阵 U 所展开的子空间中的特征表达 V ,计算步骤如下:

i) 随机初始化 U , V 和 E 矩阵,令中间变量 $X_e = X - E$;

ii) 固定 E ,迭代计算 U 和 V ,计算公式如下:

$$U_{ij} \leftarrow U_{ij} \frac{(\tilde{X}_e \tilde{V}^T)_{ij}}{(\tilde{U} \tilde{V}^T)_{ij}} \quad 4$$

$$V_{ij} \leftarrow V_{ij} \sqrt{\frac{(U^T X_e + \lambda_3 V G^-)_{ij}}{(U^T U V + \lambda_3 V G^+)_{ij}}} \quad 5$$

其中 $\tilde{X}_e = (X_e, 0_{d \times r})$, $\tilde{V} = (V, \sqrt{\lambda_2} I_r)$, $G^+ = \frac{1}{2}(G + |G|)$, $G^- = \frac{1}{2}(G - |G|)$, $|G|$ 为矩阵 G 中所有元素取对应绝对值所构成的矩阵;

iii) 固定 U 和 V ,计算最佳的 E ,计算公式如下:

$$E = f_{\frac{\lambda_1}{2}}(X - UV) \quad 6$$

其中函数 f 定义为：

$$f_v(x) = \begin{cases} x-v, & \text{若 } x > v, \\ x+v, & \text{若 } x < v, \\ 0 & \text{其他} \end{cases} \quad 7$$

4) 在进行图像检索时,将作为检索参考的图像 x 转换到图像基矩阵 U 所展开的子空间中,得到新的特征表达 v ,计算如下:

$$v = (U^T U)^{-1} U^T x = Mx, \quad 8$$

其中, $M = (U^T U)^{-1} U^T$, U^T 为矩阵 U 转置, $(U^T U)^{-1}$ 为 $U^T U$ 的逆矩阵;

5) 依据步骤 4) 得到的 v ,计算 v 与图像数据集中任一图像 x_i 的子空间特征表达 v_i 之间的距离 s_i ,计算公式如下:

$$s_i = \exp(-\|v-v_i\|^2) \quad 9$$

最后根据距离 s_i 对图像数据集中所有图像进行排序,返回距离值最小的前 K 个图像给用户,从而实现图像检索。

一种基于鲁棒非负矩阵分解的图像检索方法

技术领域

[0001] 本发明涉及基于内容的图像检索和非负矩阵分解方法,尤其涉及一种基于鲁棒非负矩阵分解的图像检索方法。

背景技术

[0002] 近年来,随着互联网络和智能手机、数码相机的普及,大量的图像资源被源源不断地产生出来。为了对这些图像资源进行有效的组织管理和检索,各种图像检索技术也随之被提出。目前,图像检索已经成为计算机视觉,尤其是多媒体检索领域里的一个研究热点。

[0003] 除了最早的基于关键字的图像检索技术,目前绝大部分的图像检索技术都是基于内容的图像检索技术。基于内容的图像检索技术是通过从图像提取表征各种图像视觉特征的视觉特征,然后基于特定的算法计算用户提交的待检索图像与图像数据库中各个图片之间的相似性,即距离,从而依据特征距离的大小,对图像进行排序,返回给用户距离值较小的图像,实现对图像数据库中图像的检索。

[0004] 在基于内容的图像检索算法中,数据降维算法是一种广泛被使用的技术。该技术的核心思想是认为原始图像特征空间存在一定的不足,无法有效地将不同类型的图像区分开来,因此希望寻找到一个合适的特征子空间,在该子空间上不同类型的图像彼此分开,而同类型的图像彼此聚拢。这类方法包括:主成分分析(Principal Component Analysis, PCA),线性鉴别分析(Linear Discriminant Analysis, LDA),非负矩阵分解(Nonnegative Matrix Factorization, NMF)和各种流形学习算法,如多维标度分析(Multidimensional Scaling, MDS),局部线性嵌套(Local Linear Embedded, LLE),局部保持投影(Local Preserving Projection, LPP)等。

[0005] 在这些子空间方法中,NMF对降维后的数据添加了非负性约束,其目标函数是求解两个非负矩阵 U 和 V ,使得 $X \approx UV$ 。由于非负性约束的引入,使得NMF能够学习得到具有局部特征的基矩阵 U ,也获得了一种与人们认知中局部构成整体所一致的基于局部的数据表达方法。因此该方法的可解释性比较强,被广泛应用到图像检索,人脸识别,数字识别,文本分类等应用。

[0006] Liang等将NMF的初始化过程扩展为有监督的初始化,然后使用NMF在隐含语义索引中,用来寻找从底层视觉特征到高层语义特征之间的关系。BeAbdallah等人则采用NMF构建图像的隐含表达,用来提高图像的标注准确性。Caicedo等人基于NMF生成多模态图像表达,并在构建NMF目标函数的时候考虑图像语义信息。

[0007] 然而,在传统的NMF对图像数据间的关系表达和抗噪声能力方法比较弱。针对这个两方面问题,本方法借助L1范数在视觉特征的自动选择特性,构建视觉L1范数约束,同时引入稀疏噪声项约束,然后将这两种约束条件加入到传统的NMF算法框架中,得到鲁棒非负矩阵分解算法模型。基于所提出的模型,对新图像的处理计算时间为线性时间,能够快速准确地进行图像检索。

发明内容

[0008] 本发明的目的是克服现有基于非负矩阵分解在图像间关系表达和抗噪声两方面的不足,提出一种基于鲁棒非负矩阵分解的图像检索方法。

[0009] 基于鲁棒非负矩阵分解的图像检索方法包括如下步骤:

[0010] 1) 从公共图像数据集 Core15K 中选择 50 类共 5000 张图像,提取 PHOG 视觉特征 [1] 集 $X = [x_1, \dots, x_n]$, 其中 $n = 5000$, x_i 为第 i 张图像所对应的 PHOG 视觉特征向量;将 PHOG 视觉特征集 X 作为所要构建的 L1 图的顶点, L1 图的边权重矩阵 $W = [w_1, \dots, w_n]$, 其中 w_i 计算如下:

$$[0011] \quad \begin{aligned} \min & \|w_i\|_1 \\ \text{s.t.} & \|x_i - Xw_i\| \leq \varepsilon, w_{ii} = 0, \forall i \end{aligned} \quad 1$$

[0012] 其中, x_i 为第 i 幅图像的视觉特征, w_i 为矩阵 W 的第 i 列, ε 为误差系数, w_{ii} 为向量 w_i 的第 i 个分量;

[0013] 构建表征图像数据分布结构关系的 L1 图约束如下:

$$[0014] \quad \begin{aligned} \min R_g &= \sum_{i=1}^n \|v_i - Vw_i\|^2 = \|V - VW\|_F^2 \\ &= \text{tr}(V(I-W)(I-W)^T V^T) = \text{tr}(VGV^T) \end{aligned} \quad 2$$

[0015] 其中, $G = (I-W)(I-W)^T$, $V = [v_1, \dots, v_n]$ 为图像集在子空间上的特征表达, I 为单位矩阵;

[0016] 2) 将 L1 图约束和稀疏噪声约束项 E 加入到传统的非负矩阵分解算法框架中, 构建鲁棒非负矩阵分解算法模型:

$$[0017] \quad \min_{U, V \geq 0, E} \|X - UV - E\|_F^2 + \lambda_1 \|E\| + \lambda_2 \|U\|_F^2 + \lambda_3 \cdot \text{tr}(VGV^T) \quad 3$$

[0018] 其中, λ_1 , λ_2 和 λ_3 为正则化参数, 在实验中分别设置为 0.01, 0.001 和 0.1, $U \in \mathbb{R}^{d \times r}$ 为保持视觉和语义相似性的图像基矩阵, $V \in \mathbb{R}^{r \times n}$ 为图像基矩阵 U 所展开的子空间中的特征表达, $E \in \mathbb{R}^{d \times n}$ 是稀疏噪声约束项, d 为图像特征的维度, r 为低维特征空间的维度;

[0019] 3) 利用图像数据集对公式 3 算法模型进行训练, 采用迭代优化求解, 计算得到保持视觉和语义相似性的图像基矩阵 U 和所有图像在图像基矩阵 U 所展开的子空间中的特征表达 V , 计算步骤如下:

[0020] i) 随机初始化 U , V 和 E 矩阵, 令中间变量 $X_e = X - E$;

[0021] ii) 固定 E , 迭代计算 U 和 V , 计算公式如下:

$$[0022] \quad U_{ij} \leftarrow U_{ij} \frac{(\tilde{X}_e \tilde{V}^T)_{ij}}{(UVV^T)_{ij}} \quad 4$$

$$[0023] \quad V_{ij} \leftarrow V_{ij} \sqrt{\frac{(U^T X_e + \lambda_3 V G^-)_{ij}}{(U^T UV + \lambda_3 V G^+)_{ij}}} \quad 5$$

[0024] 其中 $\tilde{X}_e = (X_e, 0_{d \times r})$, $\tilde{V} = (V, \sqrt{\lambda_2} I_r)$, $G^+ = \frac{1}{2}(G + |G|)$, $G^- = \frac{1}{2}(G - |G|)$, $|G|$ 为矩阵

G 中所有元素取对应绝对值所构成的矩阵；

[0025] iii) 固定 U 和 V, 计算最佳的 E, 计算公式如下：

$$[0026] \quad E = f_{\frac{\lambda}{2}}(X - UV) \quad 6$$

[0027] 其中函数 f 定义为：

[0028]

$$f_v(x) = \begin{cases} x - v, & \text{若 } x > v, \\ x + v, & \text{若 } x < v, \\ 0 & \text{其他} \end{cases} \quad 7$$

[0029] 4) 在进行图像检索时, 将作为检索参考的图像 x 转换到图像基矩阵 U 所展开的子空间中, 得到新的特征表达 v, 计算如下：

$$[0030] \quad v = (U^T U)^{-1} U^T x = Mx, \quad 8$$

[0031] 其中, $M = (U^T U)^{-1} U^T$, U^T 为矩阵 U 转置, $(U^T U)^{-1}$ 为 $U^T U$ 的逆矩阵；

[0032] 5) 依据步骤 4) 得到的 v, 计算 v 与图像数据集中任一图像 x_i 的子空间特征表达 v_i 之间的距离 s_i , 计算公式如下：

$$[0033] \quad s_i = \exp(-||v - v_i||^2) \quad 9$$

[0034] 最后根据距离 s_i 对图像数据集中所有图像进行排序, 返回距离值最小的前 K 个图像给用户, 从而实现图像检索。

[0035] 本方法能够表征图像之间的空间结构关系, 具有较强的抗噪声能力, 在得到图像基矩阵 U 后, 对新图像的低维子空间特征表达的计算时间为线性时间, 与现有方法相比能够比较快速准确地进行图像检索。

附图说明

[0036] 图 1 是 Core15K 图像数据集上的部分样例图像；

[0037] 图 2(a) 是羚羊图像作为检索例子时, 算法检索样例图；

[0038] 图 2(b) 是羚羊图像作为检索例子时, 算法前 10 个检索结果图；

[0039] 图 3(a) 是热气球图像作为检索例子时, 算法检索样例图；

[0040] 图 3(b) 是热气球图像作为检索例子时, 算法前 10 个检索结果图；

[0041] 图 4(a) 是柜子图像作为检索例子时, 算法检索样例图；

[0042] 图 4(b) 是柜子图像作为检索例子时, 算法前 10 个检索结果图；

[0043] 图 5(a) 是公交车图像作为检索例子时, 算法检索样例图；

[0044] 图 5(b) 是公交车图像作为检索例子时, 算法前 10 个检索结果图。

具体实施方式

[0045] 基于鲁棒非负矩阵分解的图像检索方法包括如下步骤：

[0046] 1) 从公共图像数据集 Core15K 中选择 50 类共 5000 张图像, 提取 PHOG 视觉特征 [1] 集 $X = [x_1, \dots, x_n]$, 其中 $n = 5000$, x_i 为第 i 张图像所对应的 PHOG 视觉特征向量；将 PHOG 视觉特征集 X 作为所要构建的 L1 图的顶点, L1 图的边权重矩阵 $W = [w_1, \dots, w_n]$, 其中 w_i 计算如下：

$$[0047] \quad \begin{aligned} & \min \|w_i\|_1 \\ \text{s.t.} \quad & \|x_i - Xw_i\| \leq \varepsilon, w_{ii} = 0, \forall i \end{aligned} \quad 1$$

[0048] 其中, x_i 为第 i 幅图像的视觉特征, w_i 为矩阵 W 的第 i 列, ε 为误差系数, w_{ii} 为向量 w_i 的第 i 个分量;

[0049] 构建表征图像数据分布结构关系的 L1 图约束如下:

$$[0050] \quad \begin{aligned} \min R_g &= \sum_{i=1}^n \|v_i - Vw_i\|^2 = \|V - VW\|_F^2 \\ &= \text{tr}(V(I-W)(I-W)^T V^T) = \text{tr}(VGV^T) \end{aligned} \quad 2$$

[0051] 其中, $G = (I-W)(I-W)^T$, $V = [v_1, \dots, v_n]$ 为图像集在子空间上的特征表达, I 为单位矩阵;

[0052] 2) 将 L1 图约束和稀疏噪声约束项 E 加入到传统的非负矩阵分解算法框架中, 构建鲁棒非负矩阵分解算法模型:

$$[0053] \quad \min_{U, V \geq 0, E} \|X - UV - E\|_F^2 + \lambda_1 \cdot \|E\| + \lambda_2 \cdot \|U\|_F^2 + \lambda_3 \cdot \text{tr}(VGV^T) \quad 3$$

[0054] 其中, λ_1 , λ_2 和 λ_3 为正则化参数, 在实验中分别设置为 0.01, 0.001 和 0.1, $U \in \mathbb{R}^{d \times r}$ 为保持视觉和语义相似性的图像基矩阵, $V \in \mathbb{R}^{r \times n}$ 为图像基矩阵 U 所展开的子空间中的特征表达, $E \in \mathbb{R}^{d \times n}$ 是稀疏噪声约束项, d 为图像特征的维度, r 为低维特征空间的维度;

[0055] 3) 利用图像数据集对公式 3 算法模型进行训练, 采用迭代优化求解, 计算得到保持视觉和语义相似性的图像基矩阵 U 和所有图像在图像基矩阵 U 所展开的子空间中的特征表达 V , 计算步骤如下:

[0056] i) 随机初始化 U , V 和 E 矩阵, 令中间变量 $X_e = X - E$;

[0057] ii) 固定 E , 迭代计算 U 和 V , 计算公式如下:

$$[0058] \quad U_{ij} \leftarrow U_{ij} \frac{(X_e \tilde{V})_{ij}}{(UV\tilde{V})_{ij}} \quad 4$$

$$[0059] \quad V_{ij} \leftarrow V_{ij} \sqrt{\frac{(U^T X_e + \lambda_3 V G^-)_{ij}}{(U^T UV + \lambda_3 V G^+)_{ij}}} \quad 5$$

[0060] 其中 $\tilde{X}_e = (X_e, 0_{d \times r})$, $\tilde{V} = (V, \sqrt{\lambda_2} I_r)$, $G^+ = \frac{1}{2}(G + |G|)$, $G^- = \frac{1}{2}(G - |G|)$, $|G|$ 为矩阵 G 中所有元素取对应绝对值所构成的矩阵;

[0061] iii) 固定 U 和 V , 计算最佳的 E , 计算公式如下:

$$[0062] \quad E = \frac{f_{\lambda_1}}{2}(X - UV) \quad 6$$

[0063] 其中函数 f 定义为:

[0064]

$$f_v(x) = \begin{cases} x-v, & \text{若 } x > v, \\ x+v, & \text{若 } x < v, \\ 0 & \text{其他} \end{cases} \quad 7$$

[0065] 4) 在进行图像检索时,将作为检索参考的图像 x 转换到图像基矩阵 U 所展开的子空间中,得到新的特征表达 v ,计算如下:

$$[0066] \quad v = (U^T U)^{-1} U^T x = Mx, \quad 8$$

[0067] 其中, $M = (U^T U)^{-1} U^T$, U^T 为矩阵 U 转置, $(U^T U)^{-1}$ 为 $U^T U$ 的逆矩阵;

[0068] 5) 依据步骤 4) 得到的 v , 计算 v 与图像数据集中任一图像 x_i 的子空间特征表达 v_i 之间的距离 s_i , 计算公式如下:

$$[0069] \quad s_i = \exp(-\|v-v_i\|^2) \quad 9$$

[0070] 最后根据距离 s_i 对图像数据集中所有图像进行排序,返回距离值最小的前 K 个图像给用户,从而实现图像检索。

[0071] 参考文献

[0072] [1]PHOG 特征代码实现:<http://www.robots.ox.ac.uk/~vgg/research/caltech/phog.html>。

[0073] 实施例 1

[0074] 采用 Core15K 图像数据集(如图 1 所示)进行试验,图 2、图 3,图 4,图 5 分别展示四个检索结果。下面结合前面所述的方法步骤,具体说明该实施实例如下:

[0075] 1)对 Core15K 中的图像分别提取图像的 680 维的 PHOG 原始图像视觉特征,将其作为图像特征表达,得到图像视觉特征矩阵 $X = [x_1, \dots, x_n]$, $n = 5000$ 。

[0076] 2) 根据图像视觉特征矩阵 X , 对视觉相似性建模,以图像集原始视觉特征 $X = [x_1, \dots, x_n]$ 作为 $L1$ 图的顶点,图的边权重矩阵 $W = [w_1, \dots, w_n]$ 中 w_i 计算如下:

$$[0077] \quad \begin{aligned} & \min \|w_i\|_1 \\ & s.t. \quad \|x_i - Xw_i\| \leq \varepsilon, w_{ii} = 0, \forall i \end{aligned} \quad 1$$

[0078] 设置 $\varepsilon = 0.001$, 采用 $L1_LS$ 求解工具包对上面目标函数 1 进行优化求解,计算得到 W 矩阵,从而得到如下一个视觉 $L1$ 图约束:

$$[0079] \quad \begin{aligned} \min R_g &= \sum_{i=1}^n \|v_i - Vw_i\|^2 = \|V - VW\|_F^2 \\ &= \text{tr}(V(I-W)(I-W)^T V^T) = \text{tr}(VGV^T) \end{aligned} \quad 2$$

[0080] 其中, $G = (I-W)(I-W)^T$, $V = [v_1, \dots, v_n]$ 为图像集在子空间上的特征表达, I 为单位矩阵。

[0081] 3) 将 $L1$ 图约束和稀疏噪声约束项 E 加入到传统的非负矩阵分解算法框架中,构建鲁棒非负矩阵分解算法模型:

$$[0082] \quad \min_{U, V \geq 0, E} \|X - UV - E\|_F^2 + \lambda_1 \cdot \|E\| + \lambda_2 \cdot \|U\|_F^2 + \lambda_3 \cdot \text{tr}(VGV^T) \quad 3$$

[0083] 其中, λ_1 , λ_2 和 λ_3 为正则化参数,分别设置为 1.6, 0.1 和 0.01, $U \in R^{d \times r}$ 为保持视觉和语义相似性的图像基矩阵, $V \in R^{r \times n}$ 为图像基矩阵 U 所展开的子空间中的特征表达, $E \in R^{d \times n}$ 是稀疏噪声约束项, d 为图像特征的维度,即 680 维, r 为低维特征空间的维度,设置为 150 维;

[0084] 4) 利用图像数据集对公式 3 算法模型进行训练, 采用迭代优化求解, 计算得到保持视觉和语义相似性的图像基矩阵 U 和所有图像在图像基矩阵 U 所展开的子空间中的特征表达 V , 计算步骤如下:

[0085] i) 随机初始化 U, V 和 E 矩阵, 令中间变量 $X_e = X - E$;

[0086] ii) 固定 E , 迭代计算 U 和 V , 计算公式如下:

$$[0087] \quad U_{ij} \leftarrow U_{ij} \frac{(\tilde{X}_e \tilde{V}^T)_{ij}}{(UVV^T)_{ij}} \quad 4$$

$$[0088] \quad V_{ij} \leftarrow V_{ij} \sqrt{\frac{(U^T X_e + \lambda_3 V G^-)_{ij}}{(U^T UV + \lambda_3 V G^+)_{ij}}} \quad 5$$

[0089] 其中 $\tilde{X}_e = (X_e, \mathbf{0}_{d \times r})$, $\tilde{V} = (V, \sqrt{\lambda_2} I_r)$, $G^+ = \frac{1}{2}(G + |G|)$, $G^- = \frac{1}{2}(G - |G|)$, $|G|$ 为矩阵 G 中所有元素取对应绝对值所构成的矩阵;

[0090] iii) 固定 U 和 V , 计算最佳的 E , 计算公式如下:

$$[0091] \quad E = f_{\lambda_1} \left(\frac{X - UV}{2} \right) \quad 6$$

[0092] 其中函数 f 定义为:

[0093]

$$f_v(x) = \begin{cases} x - v, & \text{若 } x > v, \\ x + v, & \text{若 } x < v, \\ 0 & \text{其他} \end{cases} \quad 7$$

[0094] 5) 在进行图像检索时, 将作为检索参考的图像 x 转换到图像基矩阵 U 所展开的子空间中, 得到新的特征表达 v , 计算如下:

$$[0095] \quad v = (U^T U)^{-1} U^T x = Mx, \quad 8$$

[0096] 其中, $M = (U^T U)^{-1} U^T$, U^T 为矩阵 U 转置, $(U^T U)^{-1}$ 为 $U^T U$ 的逆矩阵;

[0097] 6) 依据步骤 5) 得到的 v , 计算 v 与图像数据集中任一图像 x_i 的子空间特征表达 v_i 之间的距离 s_i , 计算公式如下:

$$[0098] \quad s_i = \exp(-\|v - v_i\|^2) \quad 9$$

[0099] 最后根据距离 s_i 对 Core15K 图像数据集中所有图像进行排序, 返回距离值最小的前 10 个图像给用户, 从而实现图像检索。



图 1

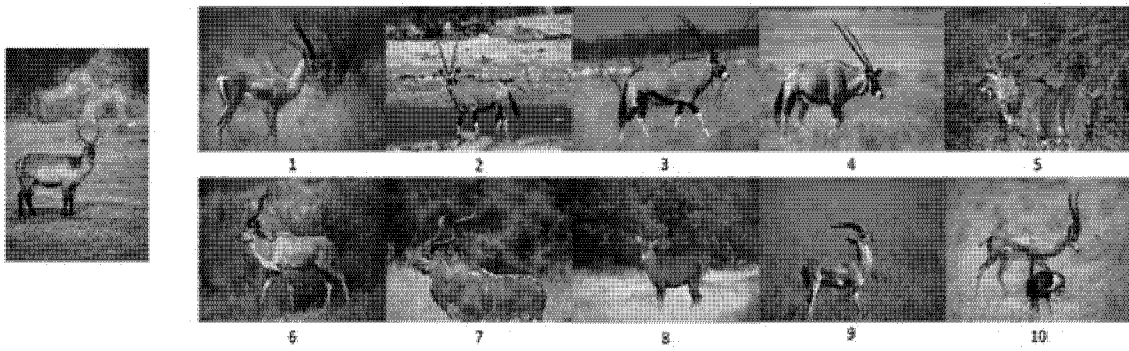


图 2 (a)

图 2 (b)

图 2

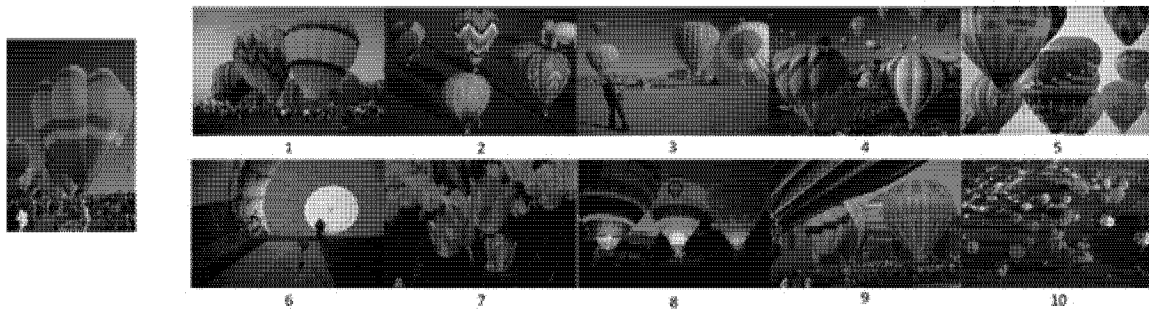


图 3 (a)

图 3 (b)

图 3

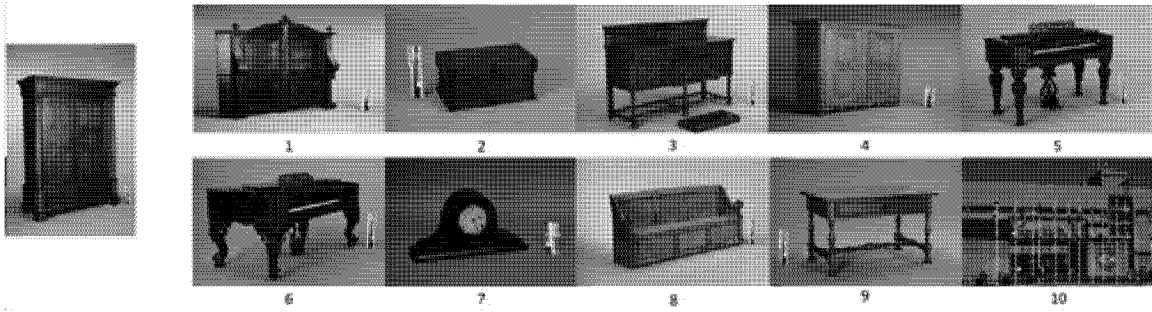


图 4(a)

图 4(b)

图 4

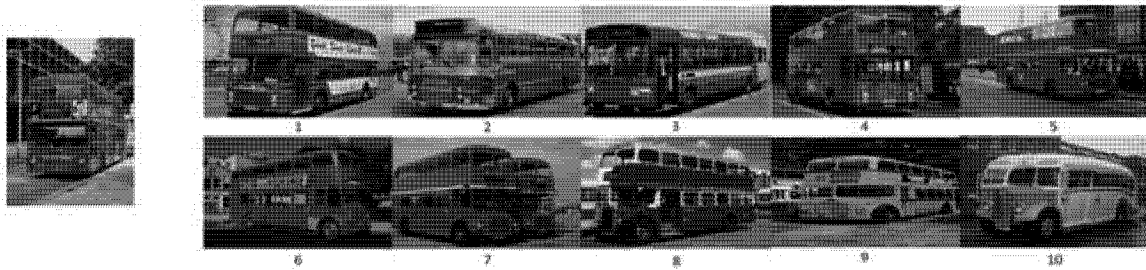


图 5(a)

图 5(b)

图 5