



# (12)发明专利

(10)授权公告号 CN 104462489 B

(45)授权公告日 2018.02.23

(21)申请号 201410800393.4

(22)申请日 2014.12.18

(65)同一申请的已公布的文献号

申请公布号 CN 104462489 A

(43)申请公布日 2015.03.25

(73)专利权人 北京邮电大学

地址 100876 北京市海淀区西土城路10号

(72)发明人 李睿凡 张光卫 鲁鹏 芦效峰

冯方向 李蕾 刘咏彬 王小捷

(74)专利代理机构 北京德琦知识产权代理有限公司

公司 11018

代理人 谢安昆 宋志强

(51)Int.Cl.

G06F 17/30(2006.01)

G06N 3/02(2006.01)

(56)对比文件

CN 102693316 A,2012.09.26,

CN 103488713 A,2014.01.01,

US 2012/0084235 A1,2012.04.05,

CN 103793507 A,2014.05.14,

审查员 王佩文

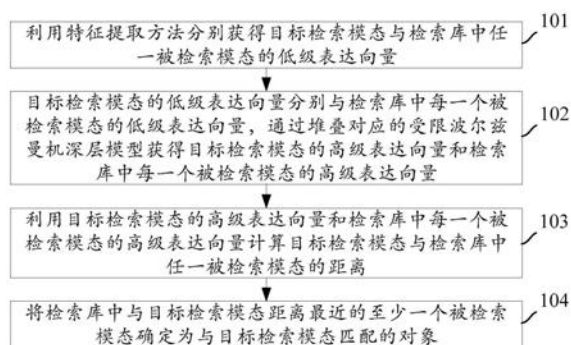
权利要求书2页 说明书6页 附图3页

(54)发明名称

一种基于深层模型的跨模态检索方法

(57)摘要

本发明提出了一种基于深层模型的跨模态检索方法,该方法包括:利用特征提取方法分别获得目标检索模态与检索库中每一个被检索模态的低级表达向量;目标检索模态的低级表达向量分别与检索库中每一个被检索模态的低级表达向量,通过堆叠对应的受限波尔兹曼机Corr-RBMs深层模型获得目标检索模态的高级表达向量和检索库中每一个被检索模态的高级表达向量;利用目标检索模态的高级表达向量和检索库中每一个被检索模态的高级表达向量计算目标检索模态与检索库中每一个被检索模态的距离;将检索库中与目标检索模态距离最近的至少一个被检索模态确定为与目标检索模态匹配的对象。



1. 一种基于深层模型的跨模态检索方法, 所述深层模型为堆叠对应的受限波尔兹曼机Corr-RBMs深层模型, 其特征在于, 该方法包括:

利用特征提取方法分别获得目标检索模态与检索库中每一个被检索模态的低级表达向量;

所述目标检索模态的低级表达向量分别与所述检索库中每一个被检索模态的低级表达向量, 通过堆叠对应的受限波尔兹曼机Corr-RBMs深层模型获得所述目标检索模态的高级表达向量和所述检索库中每一个被检索模态的高级表达向量;

利用所述目标检索模态的高级表达向量和所述检索库中每一个被检索模态的高级表达向量计算所述目标检索模态与所述检索库中每一个被检索模态的距离;

将所述检索库中与所述目标检索模态距离最近的至少一个被检索模态确定为与所述目标检索模态匹配的对象;

其中,

所述Corr-RBMs深层模型由至少两层对应的受限波尔兹曼机Corr-RBM模型堆叠而成, 所述Corr-RBMs深层模型包括第一模态Corr-RBMs和第二模态Corr-RBMs, 所述第一模态Corr-RBMs处理所述目标检索模态低级表达向量, 所述第二模态Corr-RBMs处理所述检索库中任一被检索模态的低级表达向量;

所述Corr-RBM包含第一模态受限波尔兹曼机RBM和第二模态受限波尔兹曼机RBM, 所述第一模态RBM与所述第二模态RBM包含有相同的可见层神经单元数目和相同的隐藏层神经单元数目, 所述第一模态RBM与所述第二模态RBM的隐藏层之间具有相关性约束。

2. 根据权利要求1所述的方法, 其特征在于, 该方法进一步包括:

所述Corr-RBM的配置参数  $\Theta = \{W^I, C^I, B^I, W^T, C^T, B^T\}$ , 其中, 上标I表示第一模态, 上标T表示第二模态, 具体地,  $W^I$  为第一模态RBM的各可见层神经单元与隐藏层神经单元之间的连接权值参数集合,  $C^I$  为第一模态RBM的可见层神经单元偏置参数集合,  $B^I$  为第一模态RBM的隐藏层神经单元偏置参数集合,  $W^T$  为第二模态RBM的各可见层神经单元与隐藏层神经单元之间的连接权值参数集合,  $C^T$  为第二模态RBM的可见层神经单元偏置参数集合,  $B^T$  为第二模态RBM的隐藏层神经单元偏置参数集合;

所述对应的受限波尔兹曼机Corr-RBM的配置参数  $\Theta$  为令目标函数  $F = \ell_D + \alpha \ell_I + \beta \ell_T$  最小的配置参数, 且  $\ell_D = \sum_{i=1}^m \|f_I(v_i^I) - f_T(v_i^T)\|^2$ ,  $\ell_I = -\sum_{i=1}^m \log p_I(v_i^I)$ ,  $\ell_T = -\sum_{i=1}^m \log p_T(v_i^T)$ ;

其中,  $\ell_D$  为第一模态与第二模态在嵌套空间上的距离,  $\ell_I$  为第一模态的负对数似然函数,  $\ell_T$  为第二模态的负对数似然函数;  $\alpha$  和  $\beta$  是常数, 且  $\alpha \in (0, 1)$ ,  $\beta \in (0, 1)$ ;  $f_I(\cdot)$  是第一模态RBM可见层到隐藏层的映射函数,  $f_T(\cdot)$  和第二模态RBM可见层到隐藏层的映射函数;  $p_I(\cdot)$  为第一模态RBM可见层和隐藏层神经单元的联合概率分布,  $p_T(\cdot)$  为第二模态RBM可见层和隐藏层神经单元的联合概率分布;  $\|\cdot\|$  为二范数映射;  $v$  是指RBM中的可见单元, 对应于可见变量;  $m$  是模态样本数量。

3. 根据权利要求2所述的方法, 其特征在于, 根据目标函数F确定  $\Theta$  的算法为:

A、第一模态RBM的可见层神经单元与隐藏层神经单元之间的连接权值参数集合  $w_{ij}^I$ 、可见层神经单元  $v_i^I$  的偏置  $b_i^I$  和隐藏层神经单元  $h_j^I$  的偏置  $c_j^I$  用  $\theta^I$  统一表示, 根据公式  $\theta^I \leftarrow \theta^I + \tau \cdot$

$\alpha \cdot \Delta \theta^I$ 进行更新,其中 $\tau$ 为学习速率,且 $\tau \in (0, 1)$ ;  $\alpha \in (0, 1)$ ;  $\Delta \theta^I = \{\Delta \omega_{ij}^I, \Delta b_i^I, \Delta c_j^I\}$ ,并且,

$$\Delta \omega_{ij}^I = \tau \cdot \frac{\partial \log p(v_i^I)}{\partial \omega_{ij}^I} = \tau \cdot (\langle v_i^I \cdot h_j^I \rangle_{data} - \langle v_i^I \cdot h_j^I \rangle_{model});$$

$$\Delta b_i^I = \tau \cdot \frac{\partial \log p(v_i^I)}{\partial h_j^I} = \tau \cdot (\langle h_j^I \rangle_{data} - \langle h_j^I \rangle_{model});$$

$$\Delta c_j^I = \tau \cdot \frac{\partial \log p(v_i^I)}{\partial c_j^I} = \tau \cdot (\langle v_i^I \rangle_{data} - \langle v_i^I \rangle_{model});$$

其中, $\langle \cdot \rangle_{data}$ 为经验分布下的数学期望, $\langle \cdot \rangle_{model}$ 为模型分布下的数学期望;

B、第二模态RBM的可见层神经单元与隐藏层神经单元之间的连接权值参数集合 $w_{ij}^T$ 、可见层神经单元 $v_i^T$ 的偏置 $b_i^T$ 和隐藏层神经单元 $h_j^T$ 的偏置 $c_j^T$ 用 $\theta^T$ 统一表示,根据公式 $\theta^T \leftarrow \theta^T + \tau \cdot \beta \cdot \Delta \theta^T$ 进行更新,其中, $\beta \in (0, 1)$ ;  $\Delta \theta^T = \{\Delta \omega_{ij}^T, \Delta b_i^T, \Delta c_j^T\}$ ,并且,

$$\Delta \omega_{ij}^T = \tau \cdot \frac{\partial \log p(v_i^T)}{\partial \omega_{ij}^T} = \tau \cdot (\langle v_i^T \cdot h_j^T \rangle_{data} - \langle v_i^T \cdot h_j^T \rangle_{model});$$

$$\Delta b_i^T = \tau \cdot \frac{\partial \log p(v_i^T)}{\partial h_j^T} = \tau \cdot (\langle h_j^T \rangle_{data} - \langle h_j^T \rangle_{model});$$

$$\Delta c_j^T = \tau \cdot \frac{\partial \log p(v_i^T)}{\partial c_j^T} = \tau \cdot (\langle v_i^T \rangle_{data} - \langle v_i^T \rangle_{model});$$

C、根据以下公式使用梯度下降的方法更新 $\ell_D$ :

$$\omega_{ij}^I \leftarrow \omega_{ij}^I - \tau \cdot (h_j^T - h_j^I) \cdot \delta'(h_j^I) \cdot v_i^I;$$

$$b_i^I \leftarrow b_i^I - \tau \cdot (h_j^T - h_j^I) \cdot \delta'(h_j^I);$$

$$\omega_{ij}^T \leftarrow \omega_{ij}^T - \tau \cdot (h_j^I - h_j^T) \cdot \delta'(h_j^T) \cdot v_i^T;$$

$$b_i^T \leftarrow b_i^T - \tau \cdot (h_j^I - h_j^T) \cdot \delta'(h_j^T);$$

其中, $\delta'(\cdot) = \delta(\cdot) (1 - \delta(\cdot))$ ,且 $\delta(\cdot)$ 为Logistic激活函数 $\delta(x) = 1 / (1 + \exp(-x))$ ; 重复步骤A~C,直至该算法收敛。

## 一种基于深层模型的跨模态检索方法

### 技术领域

[0001] 本发明涉及多媒体检索技术,特别是一种基于深层模型的跨模态检索方法。

### 背景技术

[0002] 近些年互联网的发展使得多模态的数据呈现爆炸式增长。例如,电子商务网站上的产品通常包含主干文字、简短的文本描述、以及相关的图片;社交网站上分享的图片通常伴有标记的描述词;一些在线新闻上包含的图片和视频信息比单纯的文字报道更具有吸引力,多模态数据的快速增长带来了巨大的跨模态检索需求。

[0003] 与传统的单模态检索不同,跨模态检索更多关注不同模态间的关系。因此,跨模态检索问题包含两个挑战问题:一是来自不同模态的数据具有完全不同的统计特性,这使得很难直接获得不同模态数据的关联关系;二是从不同模态数据中抽取的特征通常具有高维的特性并且数据集的规模非常大,这使得高效的检索不容易实现。

### 发明内容

[0004] 有鉴于此,本发明提供了一种基于深层模型的跨模态检索方法,应用深层模型解决跨模态数据的处理问题,使得经深层模型处理后的跨模态数据能够高效的进行距离计算,从而得到较优的检索结果。本发明提出的技术方案是:

[0005] 一种基于深层模型的跨模态检索方法,该方法包括:

[0006] 利用特征提取方法分别获得目标检索模态与检索库中每一个被检索模态的低级表达向量;

[0007] 所述目标检索模态的低级表达向量分别与所述检索库中每一个被检索模态的低级表达向量,通过堆叠对应的受限波尔兹曼机Corr-RBMs深层模型获得所述目标检索模态的高级表达向量和所述检索库中每一个被检索模态的高级表达向量;

[0008] 利用所述目标检索模态的高级表达向量和所述检索库中每一个被检索模态的高级表达向量计算所述目标检索模态与所述检索库中每一个被检索模态的距离;

[0009] 将所述检索库中与所述目标检索模态距离最近的至少一个被检索模态确定为与所述目标检索模态匹配的对象。

[0010] 综上所述,本发明技术方案提出了一种基于深层模型的跨模态检索方法,对于跨模态原始数据进行特征提取获得的低级表达,通过堆叠对应的受限波尔兹曼机(Corr-RBM, Correspondence Restricted Boltzmann Machine)的Corr-RBMs深层模型的处理,得到跨模态数据在相同表示空间中的低维高级表达,进而对跨模态数据的低维高级表达进行距离计算,根据距离确定检索结果。

### 附图说明

[0011] 图1为本发明技术方案的流程图;

[0012] 图2为本发明Corr-RBMs深层模型神经网络结构图;

- [0013] 图3为本发明Corr-RBM模型神经网络结构图；  
[0014] 图4为受限波尔兹曼机RBM模型的结构图；  
[0015] 图5为根据目标函数F确定 $\Theta$ 的方法流程图；  
[0016] 图6为本发明实施例的流程图。

### 具体实施方式

[0017] 为解决跨模态间的检索问题,本发明提出一种基于Corr-RBMs深层模型的跨模态检索方法,本发明技术方案的流程图如图1所示,包括以下步骤:

[0018] 步骤101:利用特征提取方法分别获得目标检索模态与检索库中任一被检索模态的低级表达向量。

[0019] 本步骤中,为在检索库中检索与目标检索模态匹配的对象,首先需要对目标检索模态和检索库中任一被检索模态的低级表达向量,特征提取方法获得的低级表达向量一般维数较高,且不同模态的低级表达向量元素各异,一般不能直接用于检索运算。

[0020] 步骤102:目标检索模态的低级表达向量分别与检索库中每一个被检索模态的低级表达向量,通过堆叠对应的受限波尔兹曼机Corr-RBMs深层模型获得目标检索模态的高级表达向量和检索库中每一个被检索模态的高级表达向量。

[0021] 本步骤中,将目标检索模态的低级表达向量分别与检索库中每一个被检索模态的低级表达向量作为一个组合,通过堆叠对应的受限波尔兹曼机Corr-RBMs深层模型获得目标检索模态的高级表达向量和检索库中每一个被检索模态的高级表达向量。通过Corr-RBMs深层模型得到的目标检索模态的高级表达向量和检索库中每一个被检索模态的高级表达向量具有低维、空间元素一致等特点,能够高效的进行检索运算。

[0022] 步骤103:利用目标检索模态的高级表达向量和检索库中每一个被检索模态的高级表达向量计算目标检索模态与检索库中任一被检索模态的距离。

[0023] 具体地,可以用欧氏距离表示目标检索模态与检索库中每一个被检索模态的距离。

[0024] 步骤104:将检索库中与目标检索模态距离最近的至少一个被检索模态确定为与目标检索模态匹配的对象。

[0025] 本步骤中,将检索库中每个被检索模态与目标检索模态的距离进行排序,选择距离目标检索模态最近的至少一个被检索模态确定为与目标检索模态匹配的对象。

[0026] 本发明提出了一种使用堆叠Corr-RBM的Corr-RBMs深层模型进行跨模态检索的方法,图2为本发明堆叠Corr-RBM的Corr-RBMs深层模型神经网络结构图,如图2所示,Corr-RBMs深层模型由至少两层Corr-RBM模型堆叠而成,该Corr-RBMs深层模型能够由两种不同模态原始数据的低级表达获得该两种不同模态原始数据的高级表达;每层Corr-RBM模型神经网络结构图如图3所示,Corr-RBM模型是在受限波尔兹曼机RBM的基础上建立的,图4为受限波尔兹曼机的神经网络结构图,下面分别对RBM模型、Corr-RBM模型以及Corr-RBMs深层模型进行详细介绍。

[0027] (一)RBM模型:

[0028] 图4为RBM的神经网络结构图,如图4所示,RBM可见层V包含m个神经单元 $v_1 \sim v_m$ ,每个神经单元 $v_i$ 的偏置为 $b_i$ ,可见层神经单元之间没有连接;隐藏层H包含s个神经单元 $h_1 \sim$

$h_s$ , 每个神经元  $h_j$  的偏置为  $c_j$ , 可见层神经元之间没有连接; 可见层神经元  $v_i$  与隐藏层神经元  $h_j$  的连接权值为  $w_{ij}$ 。为了便于理解, 图4中仅画出了部分可见层神经元与隐藏层神经元的连接权值。

[0029] RBM具有无向图的结构, 具有Logistic激活函数  $\delta(x) = 1/(1+\exp(-x))$ , 则可见层V和隐藏层H神经元的联合概率分布为:

$$[0030] \quad p(v, h) = \frac{1}{Z} \exp(-E(v, h))$$

[0031] 其中,  $Z$  为归一化常数,  $E(v, h)$  是由RBM的可见层神经元、隐藏层神经元的不同配置定义的能量函数, 根据可见层神经元、隐藏层神经元的不同配置,  $E(v, h)$  有不同的表示, 即只要RBM的可见层神经元配置与隐藏层神经元配置确定, 就有相应的能量函数, 在此不作详细介绍。

[0032] RBM的可见层神经元  $v_i$  的偏置  $b_i$ 、隐藏层神经元  $h_j$  的偏置  $c_j$ 、可见层神经元  $v_i$  与隐藏层神经元  $h_j$  的连接权值  $w_{ij}$  的学习可以通过比照散度估计算法得到, 比照散度估计算法为较为成熟的现有技术, 在此不作详细介绍。

[0033] (二) 对应的受限波尔兹曼机Corr-RBM模型:

[0034] 图3为本发明Corr-RBM模型的结构图, 如图3所示, Corr-RBM模型包含第一模态RBM和第二模态RBM, 第一模态RBM与第二模态RBM包含有相同的可见层神经元数目  $m$  和相同的隐藏层神经元数目  $s$ , 并且第一模态RBM与第二模态RBM的隐藏层之间具有相关性约束。

[0035] 假定  $\Theta$  表示Corr-RBM模型的参数集合, 即  $\Theta = \{W^I, C^I, B^I, W^T, C^T, B^T\}$ , 其中, 上标I表示第一模态, 上标T表示第二模态, 具体地,  $W^I$  为第一模态RBM的各可见层神经元与隐藏层神经元之间的连接权值参数集合,  $C^I$  为第一模态RBM的可见层神经元偏置参数集合,  $B^I$  为第一模态RBM的隐藏层神经元偏置参数集合,  $W^T$  为第二模态RBM的各可见层神经元与隐藏层神经元之间的连接权值参数集合,  $C^T$  为第二模态RBM的可见层神经元偏置参数集合,  $B^T$  为第二模态RBM的隐藏层神经元偏置参数集合。

[0036] Corr-RBM模型的参数集合  $\Theta$  通过下面的参数学习算法进行确定:

[0037] 根据下述原则定义目标函数  $F$ : Corr-RBM模型的参数集合  $\Theta$  能够最小化第一模态与第二模态在共享表示空间上的距离, 以及最小化第一模态和第二模态的负对数似然函数。目标函数  $F$  为  $F = l_D + \alpha l_I + \beta l_T$ , 即  $\Theta$  为令  $F$  最小的参数集合。

[0038] 其中,

$$[0039] \quad \ell_D = \sum_{i=1}^m \|f_I(v_i^I) - f_T(v_i^T)\|^2;$$

$$[0040] \quad \ell_I = -\sum_{i=1}^m \log p_I(v_i^I);$$

$$[0041] \quad \ell_T = -\sum_{i=1}^m \log p_T(v_i^T)。$$

[0042] 其中,  $l_D$  为第一模态与第二模态在嵌套空间上的距离,  $l_I$  为第一模态的负对数似然函数,  $l_T$  为第二模态的负对数似然函数,  $\alpha$  和  $\beta$  是常数,  $\alpha \in (0, 1)$ ,  $\beta \in (0, 1)$ ;  $f_I(\cdot)$  是第一模态RBM可见层到隐藏层的映射函数,  $f_T(\cdot)$  是第二模态RBM可见层到隐藏层的映射函数;  $p_I$

( $\bullet$ ) 为第一模态RBM可见层和隐藏层神经单元的联合概率分布,  $p_T(\bullet)$  为第二模态RBM可见层和隐藏层神经单元的联合概率分布,  $||\bullet||$  为二范数映射。

[0043] 为了根据目标函数F确定 $\Theta$ , 可以采用交替迭代的优化过程, 首先对两个似然函数 $l_I$ 和 $l_T$ 采用比照散度估记算法进行更新, 然后使用梯度下降法更新 $l_D$ , 收敛性可以在验证集上使用跨模态检索进行检测, 具体地, 图5为根据目标函数F确定 $\Theta$ 的流程图, 包括以下步骤:

[0044] 步骤501: 利用比照散度估记算法更新第一模态RBM的参数。

[0045] 第一模态RBM的可见层神经单元与隐藏层神经单元之间的连接权值参数集合 $w_{ij}^I$ 、可见层神经单元 $v_i^I$ 的偏置 $b_i^I$ 和隐藏层神经单元 $h_j^I$ 的偏置 $c_j^I$ 用 $\theta^I$ 统一表示, 根据公式 $\theta^I \leftarrow \theta^I + \tau \cdot \alpha \cdot \Delta \theta^I$ 进行更新, 其中 $\tau$ 为学习速率,  $\tau \in (0, 1)$ ;  $\alpha \in (0, 1)$ ;  $\Delta \theta^I = \{\Delta \omega_{ij}^I, \Delta b_i^I, \Delta c_j^I\}$ , 并且,

$$[0046] \quad \Delta \omega_{ij}^I = \tau \cdot \frac{\partial \log p(v_i^I)}{\partial \omega_{ij}^I} = \tau \cdot (\langle v_i^I \cdot h_j^I \rangle_{data} - \langle v_i^I \cdot h_j^I \rangle_{model});$$

$$[0047] \quad \Delta b_i^I = \tau \cdot \frac{\partial \log p(v_i^I)}{\partial h_j^I} = \tau \cdot (\langle h_j^I \rangle_{data} - \langle h_j^I \rangle_{model});$$

$$[0048] \quad \Delta c_j^I = \tau \cdot \frac{\partial \log p(v_i^I)}{\partial c_j^I} = \tau \cdot (\langle v_i^I \rangle_{data} - \langle v_i^I \rangle_{model});$$

[0049] 其中,  $\langle \bullet \rangle_{data}$  为经验分布下的数学期望,  $\langle \bullet \rangle_{model}$  为模型分布下的数学期望;

[0050] 步骤502: 利用比照散度估记算法更新第二模态RBM的参数。

[0051] 第二模态RBM的可见层神经单元与隐藏层神经单元之间的连接权值参数集合 $w_{ij}^T$ 、可见层神经单元 $v_i^T$ 的偏置 $b_i^T$ 和隐藏层神经单元 $h_j^T$ 的偏置 $c_j^T$ 用 $\theta^T$ 统一表示, 根据公式 $\theta^T \leftarrow \theta^T + \tau \cdot \beta \cdot \Delta \theta^T$ 进行更新, 其中 $\beta \in (0, 1)$ ;  $\Delta \theta^T = \{\Delta \omega_{ij}^T, \Delta b_i^T, \Delta c_j^T\}$ , 并且,

$$[0052] \quad \Delta \omega_{ij}^T = \tau \cdot \frac{\partial \log p(v_i^T)}{\partial \omega_{ij}^T} = \tau \cdot (\langle v_i^T \cdot h_j^T \rangle_{data} - \langle v_i^T \cdot h_j^T \rangle_{model});$$

$$[0053] \quad \Delta b_i^T = \tau \cdot \frac{\partial \log p(v_i^T)}{\partial h_j^T} = \tau \cdot (\langle h_j^T \rangle_{data} - \langle h_j^T \rangle_{model});$$

$$[0054] \quad \Delta c_j^T = \tau \cdot \frac{\partial \log p(v_i^T)}{\partial c_j^T} = \tau \cdot (\langle v_i^T \rangle_{data} - \langle v_i^T \rangle_{model});$$

[0055] 步骤503: 使用梯度下降的方法更新第一模态与第二模态在嵌套空间上的距离。

[0056] 具体地, 根据以下公式使用梯度下降的方法更新第一模态与第二模态在嵌套空间上的距离 $l_D$ :

$$[0057] \quad \omega_{ij}^I \leftarrow \omega_{ij}^I - \tau \cdot (h_j^T - h_j^I) \cdot \delta'(h_j^I) \cdot v_i^I;$$

$$[0058] \quad b_i^I \leftarrow b_i^I - \tau \cdot (h_j^T - h_j^I) \cdot \delta'(h_j^I);$$

[0059]  $\omega_{ij}^T \leftarrow \omega_{ij}^T - \tau \cdot (h_j^I - h_j^T) \cdot \delta'(h_j^T) \cdot v_i^T;$

[0060]  $b_i^T \leftarrow b_i^T - \tau \cdot (h_j^I - h_j^T) \cdot \delta'(h_j^T);$

[0061] 其中,  $\delta'(\cdot) = \delta(\cdot)(1 - \delta(\cdot))$ , 且  $\delta(\cdot)$  为 Logistic 激活函数  $\delta(x) = 1/(1 + \exp(-x))$ 。

[0062] 步骤504: 重复步骤501~503, 直至算法收敛。

[0063] 通过上述方法即可获得 Corr-RBM 模型的参数集合  $\Theta$ 。

[0064] (三) Corr-RBMs 深层模型

[0065] 图2为 Corr-RBMs 深层模型神经网络结构图, 如图2所示, Corr-RBMs 深层模型由至少两层对应的受限波尔兹曼机 Corr-RBM 模型堆叠而成, Corr-RBMs 深层模型包括第一模态 Corr-RBMs 和第二模态 Corr-RBMs, 第一模态 Corr-RBMs 处理目标检索模态低级表达, 第二模态 Corr-RBMs 处理检索库中任一被检索模态的低级表达。

[0066] 底层 Corr-RBM 的第一模态 RBM 可见层神经单元的输入为第一模态原始数据经特征提取获得的第一模态的低级表达, 底层 Corr-RBM 的第二模态 RBM 可见层神经单元的输入为第二模态原始数据经特征提取获得的第一模态的低级表达, 由原始数据经过特定提取获得低级表达为现有技术, 在此不作详细介绍。

[0067] 顶层 Corr-RBM 的第一 RBM 隐藏层输出第一模态的高级表达, 顶层 Corr-RBM 的第二 RBM 隐藏层输出第二模态的高级表达。

[0068] 为使本发明的目的、技术方案和优点表达的更加清楚明白, 下面结合附图及具体实施例对本发明再作进一步详细的说明。

[0069] 本实施例假设检索库包括 N 个被检索模态, 以在该检索库中检索与图片 P 相关的对象为例对本发明技术方案进行说明, 图6为本实施例的流程图, 如图6所示, 包括以下步骤:

[0070] 步骤601: 采用特征提取方法获得检索库中各被检索模态的低级表达, 以及图片 P 的低级表达。

[0071] 本步骤中, 检索库中的被检索模态的模态种类不限定, 可能为图像模态, 可能为文本模态, 也可能为语音模态, 不同模态的原始数据目前均有较成熟的特征提取方法, 例如图像模态可以应用 MPEG-7 和 Gist 描述符进行特征提取, 文本模态可以应用词袋模型进行特征提取等, 在此不再对获得图片 P 与检索库中各被检索模态的低级表达的过程进行详细描述。

[0072] 步骤602: 图片 P 的低级表达分别与检索库中每个被检索模态的低级表达通过 Corr-RBMs 深层模型处理, 获得图片 P 的高级表达与检索库中每个被检索模态的高级表达, 再利用图片 P 的高级表达与检索库中每个被检索模态的高级表达进行欧氏距离计算, 计算图片 P 与检索库中每个被检索模态的欧氏距离。

[0073] 本步骤中, 将检索库中任一个被检索模态与图片 P 作为一个组合, 通过 Corr-RBMs 深层模型对组合中的被检索模态低级表达和图片 P 的低级表达进行处理, 获得该组合中被检索模态的高级表达、图片 P 的高级表达, 然后根据欧氏距离计算公式计算图片 P 与该被检索模态的欧氏距离。

[0074] 一般的, 对于 n 维欧式空间中的两点 t 和 y, 他们的距离 d 的计算公式为



$d = \sqrt{\sum_{i=1}^n (t_i - y_i)^2}$  , 以此计算图片P和任一被检索模态的欧氏距离。

[0075] 步骤603: 根据图片P与检索库中每个被检索模态的欧氏距离由低到高进行排序, 选择排在前面的K个被检索模态作为检索结果输出。

[0076] 本实施例通过Corr-RBMs深层模型对图片模态的低级表达和检索库中各被检索模态的低级表达进行处理, 获得各自的高级表达, 再利用高级表达进行欧氏距离计算能够高效的获得检索结果。

[0077] 以上所述仅为本发明的较佳实施例而已, 并不用以限制本发明, 凡在本发明的精神和原则之内所做的任何修改、等同替换、改进等, 均应包含在本发明保护的范围之内。

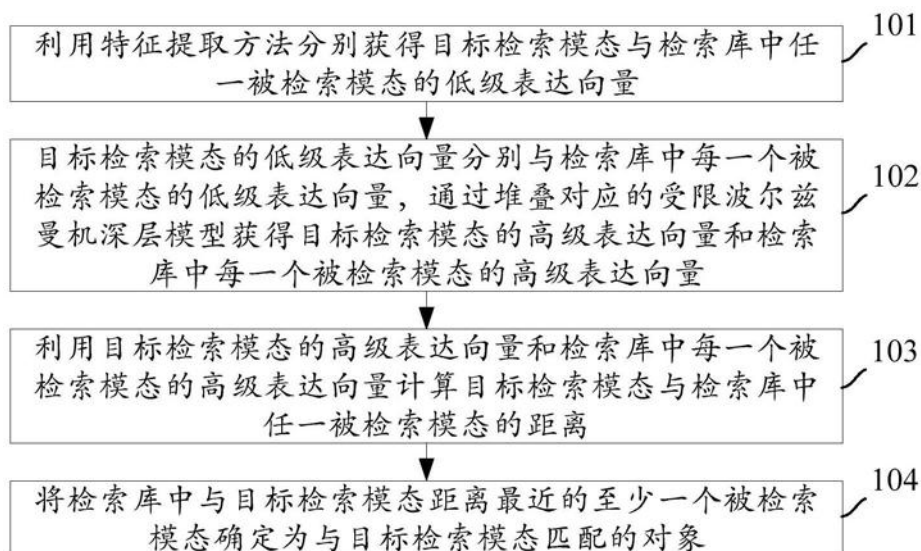


图1

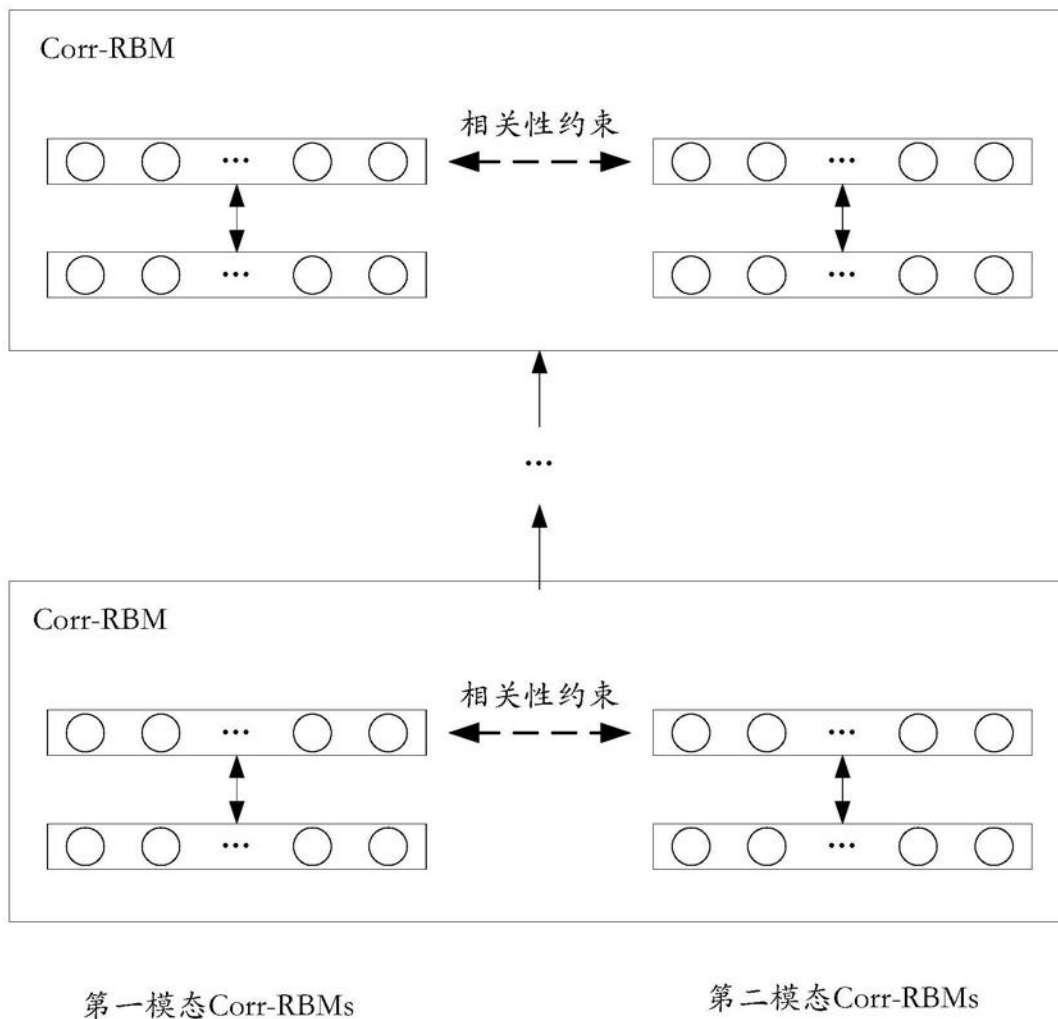


图2

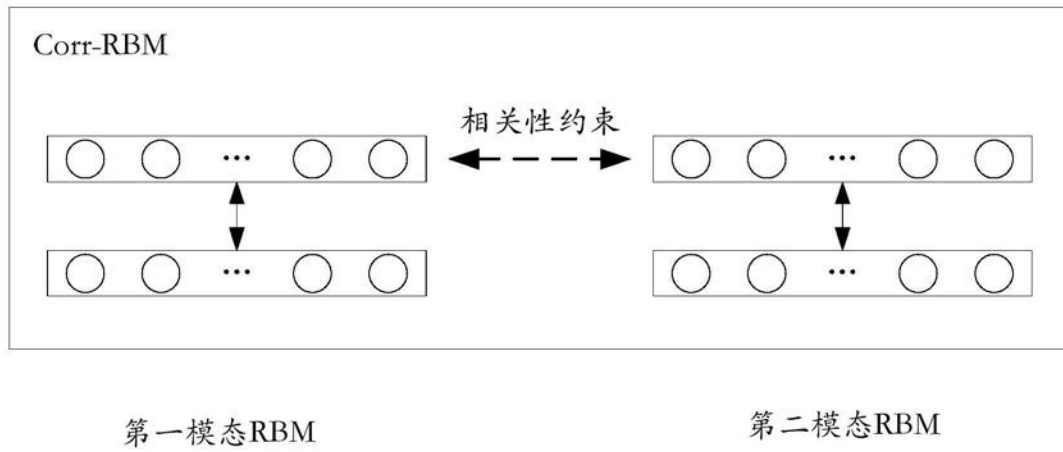


图3

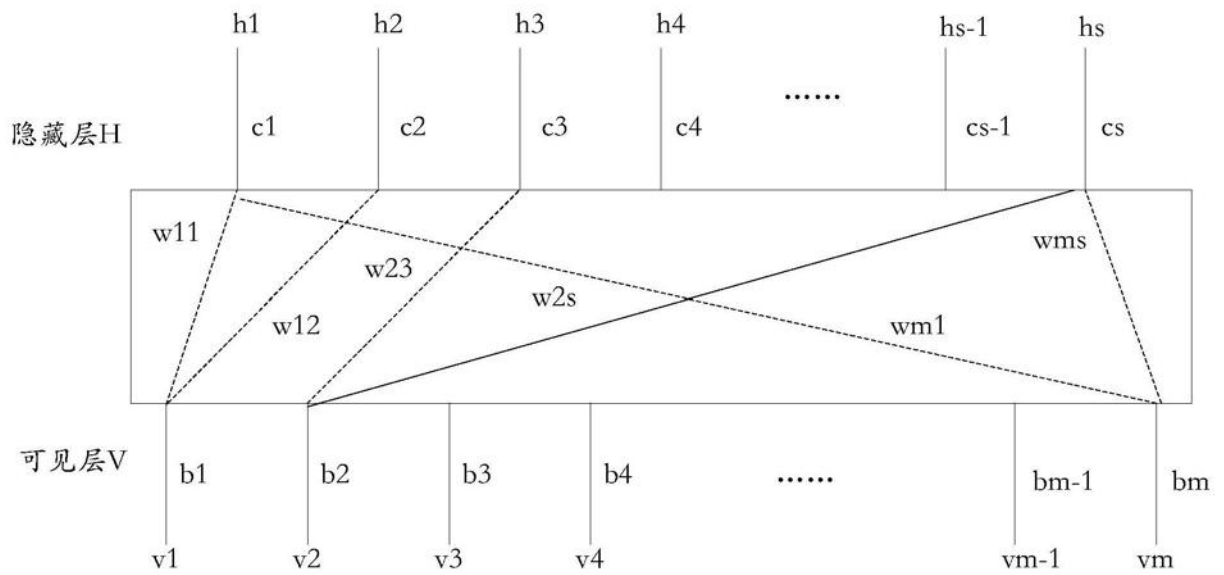


图4

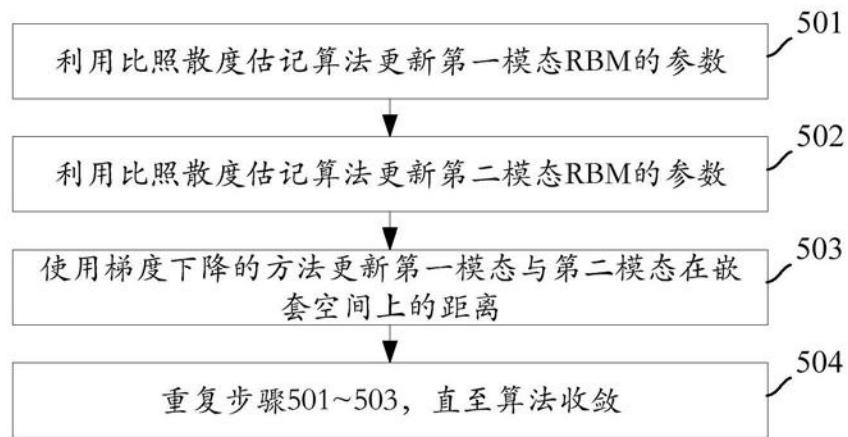


图5

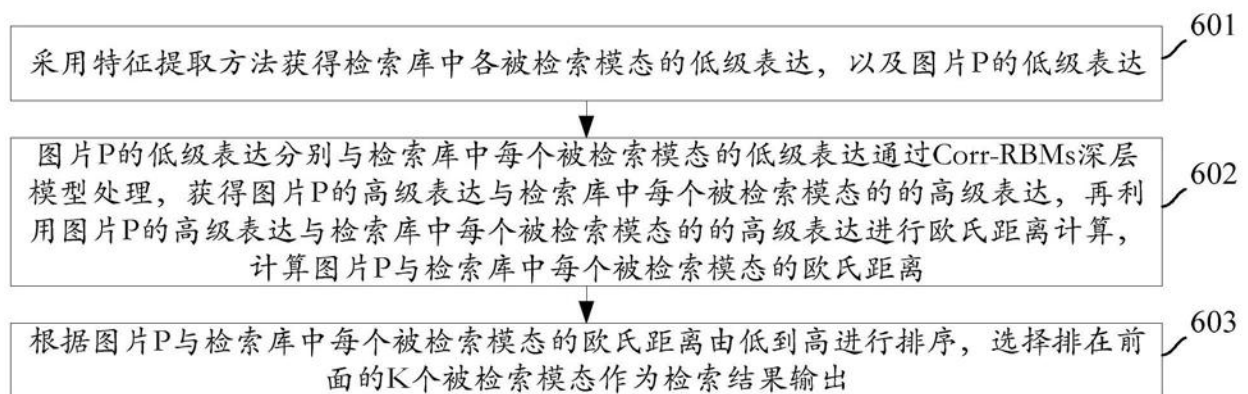


图6