



(12)发明专利

(10)授权公告号 CN 104462485 B

(45)授权公告日 2017.12.05

(21)申请号 201410797791.5

(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 104462485 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 芦效峰 鲁鹏 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)

(56)对比文件

CN 103793507 A,2014.05.14,

何宁.《图像检索中跨模语义信息获取方法研究》.《CNKI》.2013,正文第1-43页.

审查员 高丹丹

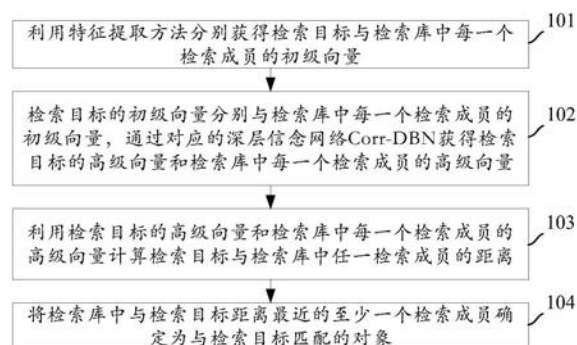
权利要求书2页 说明书6页 附图4页

(54)发明名称

一种基于对应的深层信念网络的跨模态检索方法

(57)摘要

本发明提出了一种基于对应的深层信念网络的跨模态检索方法,该方法包括:利用特征提取方法分别获得检索目标与检索库中每一个检索成员的初级向量;检索目标的初级向量分别与检索库中每一个检索成员的初级向量,通过对应的深层信念网络Corr-DBN获得检索目标的高级向量和检索库中每一个检索成员的高级向量;利用检索目标的高级向量和检索库中每一个检索成员的高级向量计算检索目标与检索库中任一检索成员的距离;将检索库中与检索目标距离最近的至少一个检索成员确定为与检索目标匹配的对象。



1. 一种基于对应的深层信念网络的跨模态检索方法,其特征在于,该方法包括:

利用特征提取方法分别获得检索目标与检索库中每一个检索成员的初级向量;

将所述检索目标和所述检索库中每一个检索成员分别作为第一模态和第二模态,将所述检索目标的初级向量分别与所述检索库中该检索成员的初级向量作为一个组合,通过对应的深层信念网络对组合中的所述检索目标的初级向量和该检索成员的初级向量进行处理,获得所述检索目标的高级向量和该检索成员的高级向量;

利用所述检索目标的高级向量和所述检索库中每一个检索成员的高级向量计算所述检索目标与所述检索库中每一个检索成员的距离;

将所述检索库中与所述检索目标距离最近的至少一个检索成员确定为与所述检索目标匹配的对象。

2. 根据权利要求1所述的方法,其特征在于,所述对应的深层信念网络Corr-DBN的非顶层为至少一层双受限波尔兹曼机RBM结构,顶层为对应的受限波尔兹曼机Corr-RBM结构,所述双RBM包括相互独立的第一模态RBM和第二模态RBM,所述Corr-RBM包括具有相关性约束的第一模态Corr-RBM和第二模态Corr-RBM。

3. 根据权利要求2所述的方法,其特征在于,所述第一模态Corr-RBM与所述第二模态Corr-RBM包含有相同的可见层神经单元数目,所述第一模态Corr-RBM与所述第二模态Corr-RBM包含有相同的隐藏层神经单元数目,所述第一模态Corr-RBM与所述第二模态Corr-RBM的隐藏层之间具有相关性约束。

4. 根据权利要求3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Corr-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Corr-RBM的可见层神经单元偏置参数集合, B^I 为第一模态Corr-RBM的隐藏层神经单元偏置参数集合, W^T 为第二模态Corr-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Corr-RBM的可见层神经单元偏置参数集合, B^T 为第二模态Corr-RBM的隐藏层神经单元偏置参数集合;

所述Corr-RBM的配置参数 Θ 为令目标函数 $Q = l_D + \alpha l_I + \beta l_T$ 最小的配置参数,且

$$\ell_D = \sum_{i=1}^m \|f_I(v_i^I) - f_T(v_i^T)\|^2, \quad \ell_I = -\sum_{i=1}^m \log p_I(v_i^I), \quad \ell_T = -\sum_{i=1}^m \log p_T(v_i^T);$$

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, α 和 β 是常数,且 $\alpha \in (0, 1)$, $\beta \in (0, 1)$; $f_I(\cdot)$ 是第一模态Corr-RBM可见层到隐藏层的映射函数, $f_T(\cdot)$ 和第二模态Corr-RBM可见层到隐藏层的映射函数; $p_I(\cdot)$ 为第一模态Corr-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Corr-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。

5. 根据权利要求4所述的方法,其特征在于,根据目标函数 Q 确定 Θ 的算法为:

A、第一模态Corr-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 θ^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 τ 为学习速率,且 $\tau \in (0, 1)$; $\alpha \in (0, 1)$; $\Delta \theta^I = \{\Delta w_{ij}^I, \Delta b_i^I, \Delta c_j^I\}$,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Corr-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 θ^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 ld :

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

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

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

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

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。

一种基于对应的深层信念网络的跨模态检索方法

技术领域

[0001] 本发明涉及多媒体检索技术,特别是一种基于对应的深层信念网络的跨模态检索方法。

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对应的深层信念网络(Correspondence Deep Belief Network, Corr-DBN)的跨模态检索方法,应用Corr-DBN解决跨模态数据的处理问题,使得经Corr-DBN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Corr-DBN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检索目标与检索库中每一个检索成员的初级向量;

[0007] 将所述检索目标和所述检索库中每一个检索成员分别作为第一模态和第二模态,将所述检索目标的初级向量分别与该检索成员的初级向量作为一个组合,通过对应的深层信念网络Corr-DBN对组合中的所述检索目标的初级向量和该检索成员的初级向量进行处理,获得所述检索目标的高级向量和该检索成员的高级向量;

[0008] 利用所述检索目标的高级向量和所述检索库中每一个检索成员的高级向量计算所述检索目标与所述检索库中每一个检索成员的距离;

[0009] 将所述检索库中与所述检索目标距离最近的至少一个检索成员确定为与所述检索目标匹配的对象。

[0010] 综上所述,本发明技术方案提出了一种基于对应的深层信念网络的跨模态检索方法,对于跨模态原始数据进行特征提取获得的初级向量,通过对应的深层信念网络Corr-DBN的处理,得到跨模态数据在相同表示空间中的低维高级向量,进而对跨模态数据的低维高级向量进行距离计算,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-DBN神经网络结构图；
[0013] 图3为本发明双受限波尔兹曼机的神经网络结构图；
[0014] 图4为受限波尔兹曼机模型的结构图；
[0015] 图5为对应的受限波尔兹曼机模型的结构图；
[0016] 图6为根据目标函数Q确定 Θ 的方法流程图；
[0017] 图7为本发明实施例的流程图。

具体实施方式

[0018] 为解决跨模态间的检索问题,本发明提出一种基于对应的深层信念网络Corr-DBN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:

[0019] 步骤101:利用特征提取方法分别获得检索目标与检索库中每一个检索成员的初级向量。

[0020] 本步骤中,为在检索库中检索与检索目标匹配的对象,首先需要对检索目标和检索库中每一检索成员进行初级向量的获取,而特征提取方法获得的初级向量一般维数较高,且不同模态的初级向量元素各异,一般不能直接用于检索运算。

[0021] 步骤102:检索目标的初级向量分别与检索库中每一个检索成员的初级向量,通过对应的深层信念网络Corr-DBN获得检索目标的高级向量和检索库中每一个检索成员的高级向量。

[0022] 本步骤中,将检索目标的初级向量分别与检索库中每一个检索成员的初级向量作为一个组合,通过对应的深层信念网络Corr-DBN获得检索目标的高级向量和检索库中每一个检索成员的高级向量。通过对应的深层信念网络Corr-DBN得到的检索目标的高级向量和检索库中每一个检索成员的高级向量具有低维、空间元素一致等特点,能够高效的直接进行检索运算。

[0023] 具体地,可以将检索目标作为第一模态,将任一检索成员作为第二模态,Corr-DBN通过处理第一模态与第二模态的初级向量,最终输出第一模态与第二模态的高级向量。

[0024] 步骤103:利用检索目标的高级向量和检索库中每一个检索成员的高级向量计算检索目标与检索库中任一检索成员的距离。

[0025] 具体地,可以用欧氏距离表示检索目标与检索库中每一个检索成员的距离。

[0026] 步骤104:将检索库中与检索目标距离最近的至少一个检索成员确定为与检索目标匹配的对象。

[0027] 本步骤中,将检索库中每个检索成员与检索目标的距离进行排序,选择距离检索目标最近的至少一个检索成员确定为与检索目标匹配的对象。

[0028] 本发明提出了一种使用对应的深层信念网络Corr-DBN进行跨模态检索的方法,图2为本发明对应的深层信念网络Corr-DBN神经网络结构图,如图2所示,Corr-DBN首先对两种模态的初级向量使用至少一层双受限波尔兹曼机(Restricted Boltzmann Machine, RBM)模型获得该两种模态的中级向量,在Corr-DBN模型的顶层通过Corr-RBM模型对两种模态的中级向量进行进一步处理,最终获得两种模态的高级向量。下面分别对双RBM模型、Corr-RBM模型以及Corr-DBN模型进行详细介绍。

[0029] (一) 双RBM模型:

[0030] 图3为双RBM的神经网络结构图,如图3所示,双RBM模型包括第一模态RBM模型和第二模态RBM模型,第一模态RBM模型和第二模态RBM模型为相互独立的两个单RBM模型,第一模态RBM模型对第一模态向量进行处理,第二模态RBM模型对第二模态向量进行处理,且第一模态RBM模型和第二模态RBM模型之间无连接。下面仅对第一模态RBM模型进行介绍,第二模态RBM模型的结构及参数设计与第一模态RBM结构相同。

[0031] 图4为第一模态RBM模型的神经网络结构图,如图4所示,第一模态RBM模型的可见层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 $h_1 \sim h_s$,每个神经单元 h_j 的偏置为 c_j ,可见层神经单元之间没有连接;可见层神经单元 v_i 与隐藏层神经单元 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单元与隐藏层神经单元的连接权值。

[0032] 单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神经单元的联合概率分布为:

$$[0033] \quad p(v, h) = \frac{1}{Z} \exp(-E(v, h))$$

[0034] 其中,Z为归一化常数, $E(v, h)$ 是由RBM的可见层神经单元、隐藏层神经单元的不同配置定义的能量函数,根据可见层神经单元、隐藏层神经单元的不同配置, $E(v, h)$ 有不同的表示,即只要RBM的可见层神经单元配置与隐藏层神经单元配置确定,就有相应的能量函数,在此不作详细介绍。

[0035] 双RBM模型的参数包括: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、可见层神经单元的偏置集合和隐藏层神经单元的偏置集合,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、可见层神经单元的偏置集合和隐藏层神经单元的偏置集合,上述参数可通过比照散度估计算法进行确定,比照散度估计算法为现有技术,在此不再详细介绍。

[0036] (二)对应的受限波尔兹曼机Corr-RBM模型:

[0037] 图5为本发明Corr-RBM模型的结构图,如图5所示,Corr-RBM模型包含第一模态Corr-RBM和第二模态Corr-RBM,第一模态Corr-RBM与第二模态Corr-RBM包含有相同的可见层神经单元数目,第一模态Corr-RBM与第二模态Corr-RBM包含有相同的隐藏层神经单元数目,并且第一模态Corr-RBM与第二模态Corr-RBM的隐藏层之间具有相关性约束。

[0038] 假定 Θ 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Corr-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Corr-RBM的可见层神经单元偏置参数集合, B^I 为第一模态Corr-RBM的隐藏层神经单元偏置参数集合, W^T 为第二模态Corr-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Corr-RBM的可见层神经单元偏置参数集合, B^T 为第二模态Corr-RBM的隐藏层神经单元偏置参数集合。

[0039] Corr-RBM模型的参数集合 Θ 通过下面的参数学习算法进行确定:

[0040] 根据下述原则定义目标函数T:Corr-RBM模型的参数集合 Θ 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Q为 $Q = l_D + \alpha l_I + \beta l_T$,即 Θ 为令Q最小的参数集合。

[0041] 其中,

$$[0042] \quad \ell_D = \sum_{i=1}^m \|f_I(v_i^I) - f_T(v_i^T)\|^2;$$

$$[0043] \quad \ell_I = -\sum_{i=1}^m \log p_I(v_i^I);$$

$$[0044] \quad \ell_T = -\sum_{i=1}^m \log p_T(v_i^T)。$$

[0045] 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, α 和 β 是常数, $\alpha \in (0, 1)$, $\beta \in (0, 1)$; $f_I(\cdot)$ 是第一模态 Corr-RBM 可见层到隐藏层的映射函数, $f_T(\cdot)$ 是第二模态 Corr-RBM 可见层到隐藏层的映射函数; $p_I(\cdot)$ 为第一模态 Corr-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 Corr-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。

[0046] 为了根据目标函数 Q 确定 Θ , 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 采用比照散度估计算法进行更新, 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6为根据目标函数 Q 确定 Θ 的流程图, 包括以下步骤:

[0047] 步骤601: 利用比照散度估计算法更新第一模态 Corr-RBM 的参数。

[0048] 第一模态 Corr-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 θ^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 τ 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,

$$[0049] \quad \Delta \omega_{ij}^I = \tau \cdot \frac{\partial \log p(v_i^I)}{\partial \omega_{ij}^I} = \tau \cdot (\langle v_i^I \cdot h_j^I \rangle_{data} - \langle v_i^I \cdot h_j^I \rangle_{model});$$

$$[0050] \quad \Delta b_i^I = \tau \cdot \frac{\partial \log p(v_i^I)}{\partial h_j^I} = \tau \cdot (\langle h_j^I \rangle_{data} - \langle h_j^I \rangle_{model});$$

$$[0051] \quad \Delta c_j^I = \tau \cdot \frac{\partial \log p(v_i^I)}{\partial c_j^I} = \tau \cdot (\langle v_i^I \rangle_{data} - \langle v_i^I \rangle_{model});$$

[0052] 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;

[0053] 步骤602: 利用比照散度估计算法更新第二模态 Corr-RBM 的参数。

[0054] 第二模态 Corr-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 θ^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,

$$[0055] \quad \Delta \omega_{ij}^T = \tau \cdot \frac{\partial \log p(v_i^T)}{\partial \omega_{ij}^T} = \tau \cdot (\langle v_i^T \cdot h_j^T \rangle_{data} - \langle v_i^T \cdot h_j^T \rangle_{model});$$

$$[0056] \quad \Delta b_i^T = \tau \cdot \frac{\partial \log p(v_i^T)}{\partial h_j^T} = \tau \cdot (\langle h_j^T \rangle_{data} - \langle h_j^T \rangle_{model});$$

$$[0057] \quad \Delta c_j^T = \tau \cdot \frac{\partial \log p(v_i^T)}{\partial c_j^T} = \tau \cdot (\langle v_i^T \rangle_{data} - \langle v_i^T \rangle_{model})$$

[0058] 步骤603: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。

[0059] 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 :

$$[0060] \quad \omega_{ij}^l \leftarrow \omega_{ij}^l - \tau \cdot (h_j^T - h_j^l) \cdot \delta'(h_j^l) \cdot v_i^l;$$

$$[0061] \quad b_i^l \leftarrow b_i^l - \tau \cdot (h_j^T - h_j^l) \cdot \delta'(h_j^l);$$

$$[0062] \quad \omega_{ij}^T \leftarrow \omega_{ij}^T - \tau \cdot (h_j^l - h_j^T) \cdot \delta'(h_j^T) \cdot v_i^T;$$

$$[0063] \quad b_i^T \leftarrow b_i^T - \tau \cdot (h_j^l - h_j^T) \cdot \delta'(h_j^T);$$

[0064] 其中, $\delta'(\cdot) = \delta(\cdot)(1 - \delta(\cdot))$, 且 $\delta(\cdot)$ 为Logistic激活函数 $\delta(x) = 1/(1 + \exp(-x))$ 。

[0065] 步骤604:重复步骤601~603,直至算法收敛。

[0066] 通过上述方法即可获得Corr-RBM模型的参数集合 Θ 。

[0067] (三)对应的深层信念网络Corr-DBN模型

[0068] 图2为Corr-DBN神经网络结构图,如图2所示,Corr-DBN的非顶层为至少一层双受限波尔兹曼机RBM结构,顶层为对应的受限波尔兹曼机Corr-RBM结构,所述双RBM包括相互独立的第一模态RBM和第二模态RBM,所述Corr-RBM包括具有相关性约束的第一模态Corr-RBM和第二模态Corr-RBM。Corr-DBN首先对两种模态的初级向量使用至少一层双RBM模型获得该两种模态的中级向量,在Corr-DBN模型的顶层通过Corr-RBM模型对两种模态的中级向量进行进一步处理,最终获得两种模态的高级向量。

[0069] 底层双RBM模型的第一模态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初级向量,底层双RBM模型的第二模态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二模态的初级向量,由原始数据经过特定提取获得初级向量为现有技术,在此不作详细介绍。

[0070] 相邻两层双RBM模型中,上层双RBM的可见层神经元个数等于下层双RBM隐藏层神经单元的个数;相邻顶层Corr-RBM的可见层神经元个数等于与该Corr-RBM相邻的双RBM模型中的隐藏层神经元个数。

[0071] 顶层Corr-RBM模型的第一Corr-RBM输出第一模态的高级向量,第二Corr-RBM输出第二模态的高级向量。

[0072] 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。

[0073] 本实施例假设检索库包括M个检索成员,以在该检索库中检索与文本J相关的对象为例对本发明技术方案进行说明,图7为本实施例的流程图,如图7所示,包括以下步骤:

[0074] 步骤701:采用特征提取方法获得检索库中各检索成员的初级向量,以及文本J的初级向量。

[0075] 本步骤中,检索库中的检索成员的模态种类不限定,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,在此不再对获得文本J以及检索库中各检索成员的初级向量的过程进行详细描述。

[0076] 步骤702:文本J的初级向量分别与检索库中每个检索成员的初级向量通过Corr-DBN模型处理,获得文本J的高级向量与检索库中每个检索成员的高级向量,再利用文本J的高级向量与检索库中每个检索成员的的高级向量进行欧氏距离计算,计算文本J与检索库中每个检索成员的欧氏距离。

[0077] 本步骤中,将检索库中任一个检索成员与文本J作为一个组合,通过Corr-DBN模型对组合中的检索成员初级向量和文本J的初级向量进行处理,获得该组合中检索成员的高级向量、文本J的高级向量,然后根据欧氏距离计算公式计算文本J与该检索成员的欧氏距离。

[0078] 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

$$d = \sqrt{\sum_{i=1}^n (t_i - y_i)^2}$$

由于经过Corr-DBN模型得到的文本J的高级向量和任一检索成员的高级

向量为同维向量,以此计算文本J和任一检索成员的欧氏距离。

[0079] 步骤703:根据文本就与检索库中每个检索成员的欧氏距离由低到高进行排序,选择排在前面的K个检索成员作为检索结果输出。

[0080] 本实施例通过Corr-DBN模型对文本模态的初级向量和检索库中各检索成员的初级向量进行处理,获得各自的高级向量,再利用高级向量进行欧氏距离计算能够高效的获得检索结果。

[0081] 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。

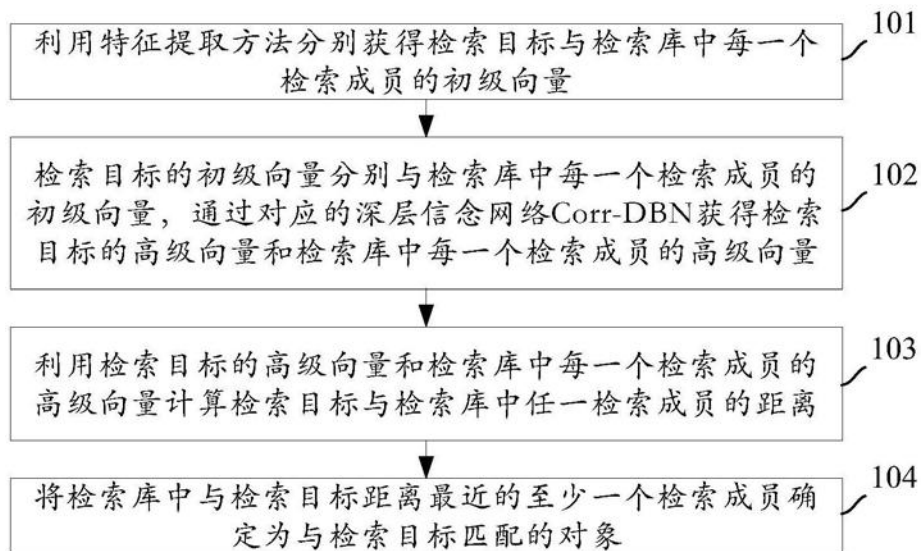


图1

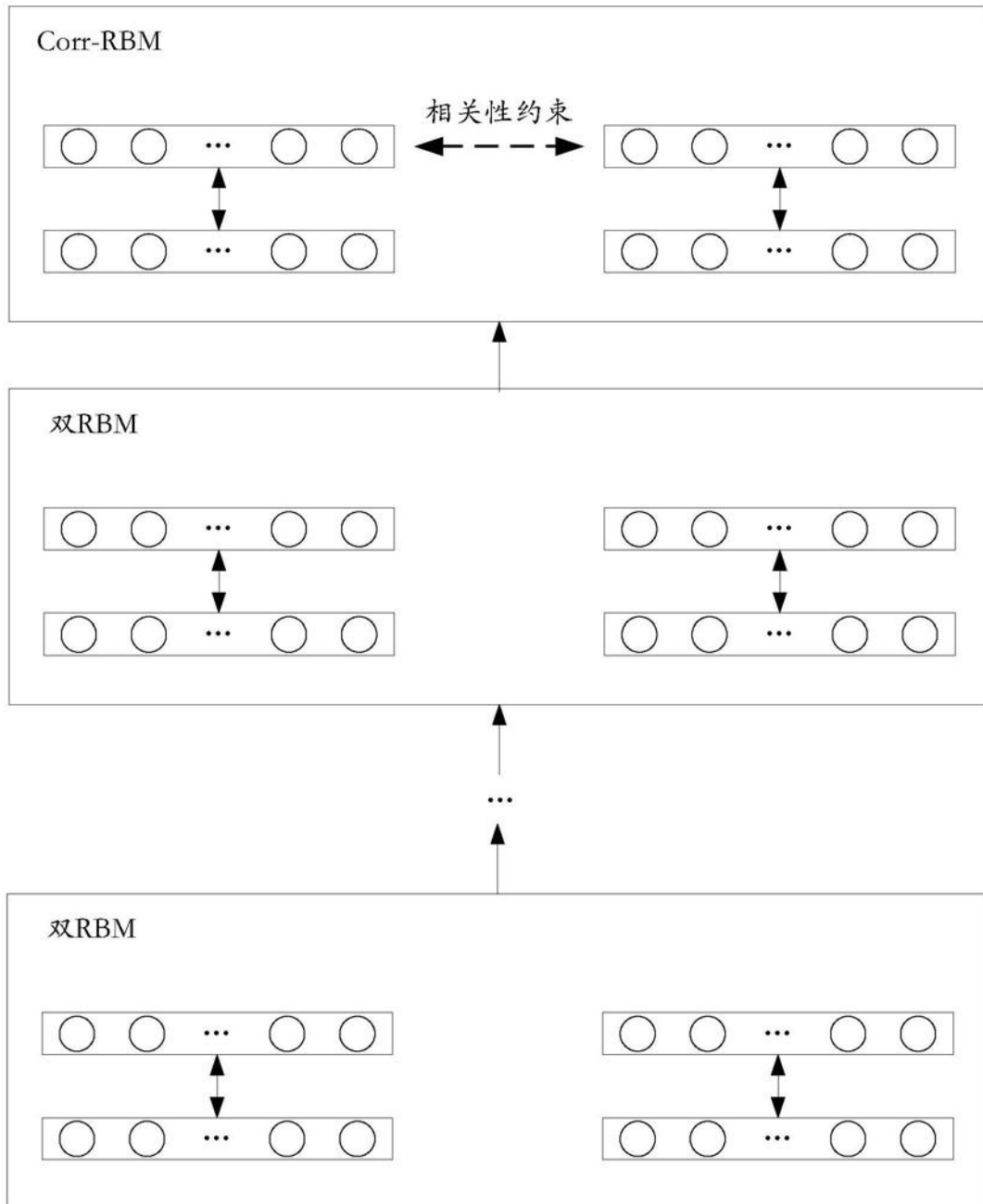


图2

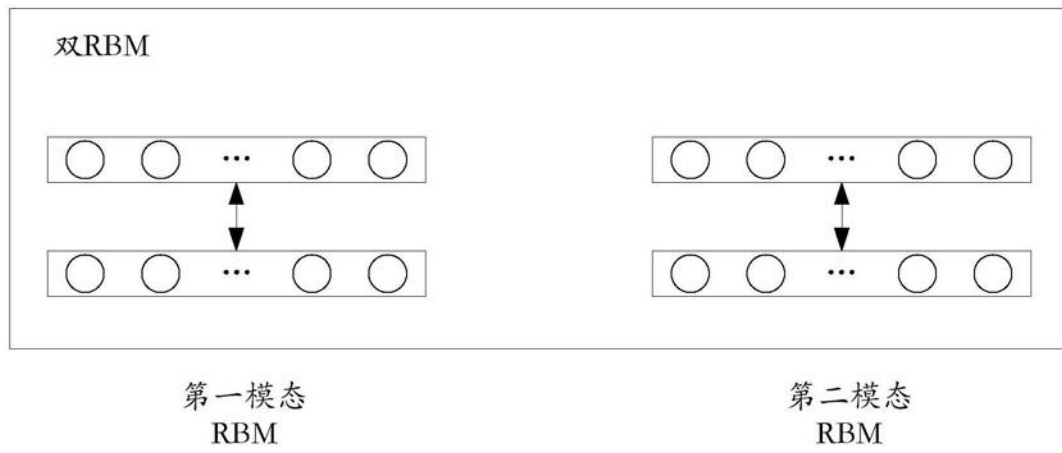


图3

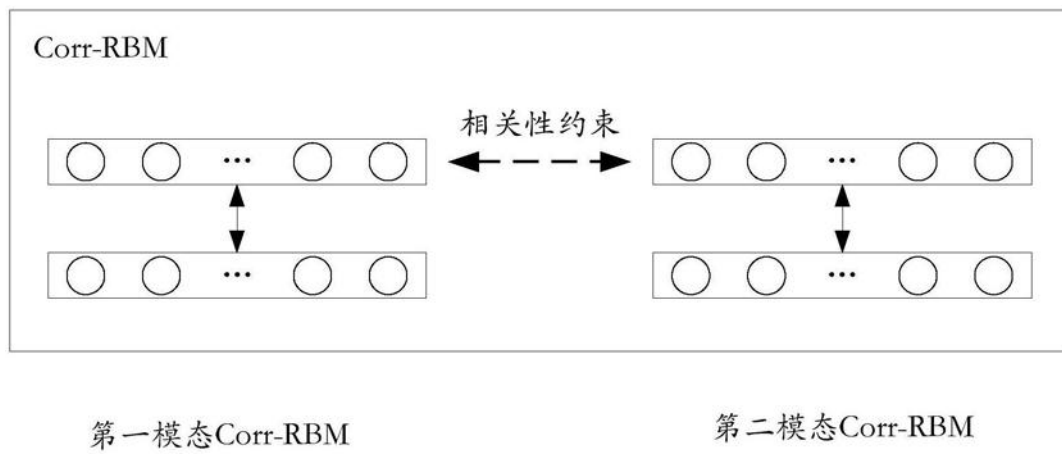


图4

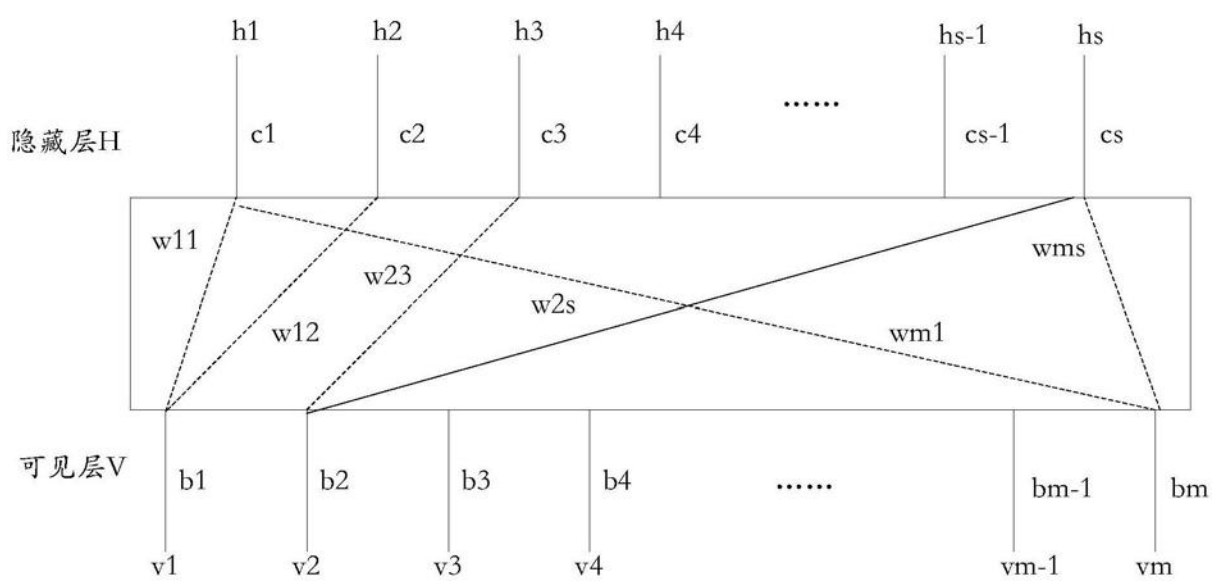


图5



图6

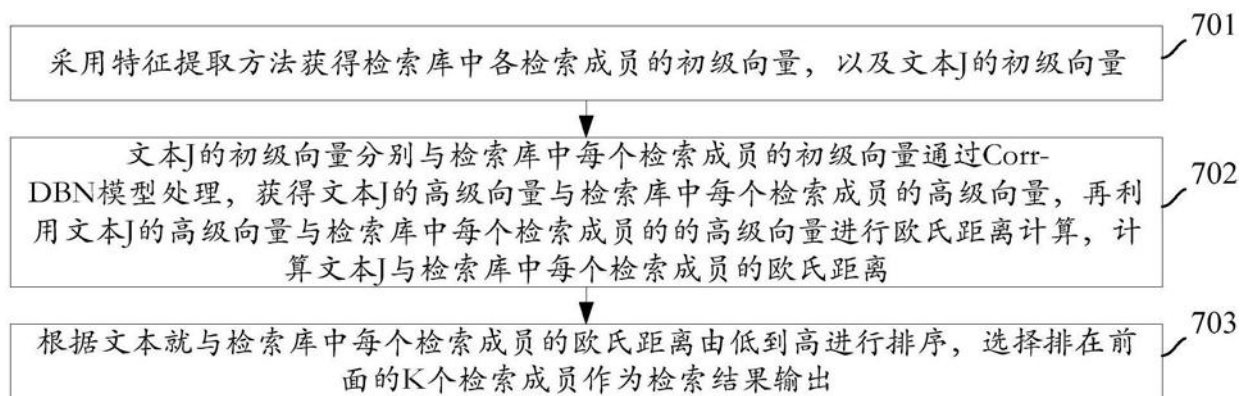


图7