



## (12)发明专利

(10)授权公告号 CN 103793507 B

(45)授权公告日 2016.10.05

(21)申请号 201410039222.4

(22)申请日 2014.01.26

(65)同一申请的已公布的文献号

申请公布号 CN 103793507 A

(43)申请公布日 2014.05.14

(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 鲁鹏 冯方向 王小捷

(74)专利代理机构 北京德琦知识产权代理有限公司 11018

代理人 郑红娟 宋志强

(51)Int.Cl.

G06F 17/30(2006.01)

G06N 3/08(2006.01)

(56)对比文件

CN 102663447 A,2012.09.12,

US 2009/0157601 A1,2009.06.18,

Jiquan Ngian et al..Multimodal deep learning.《Proceeding of the 28th International Conference on Machine Learning》.2011,689-696.

刘扬 等.基于多模态整合和时空上下文语义的跨媒体检索模型的研究.《计算机应用》.2009,第29卷(第4期),1182-1187.

审查员 游小容

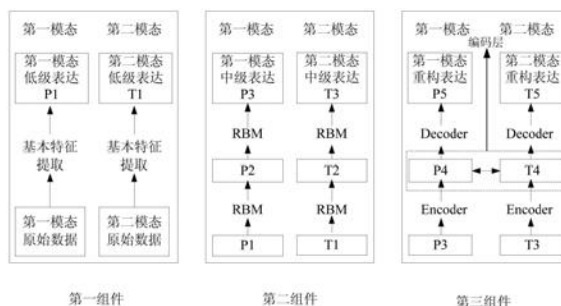
权利要求书1页 说明书6页 附图1页

(54)发明名称

一种使用深层结构获取双模态相似性测度的方法

(57)摘要

本发明提出了一种使用深层结构获取双模态相似性测度的方法,该方法包括:第一模态原始数据使用经典特征提取方法获取第一模态的低级表达P1,第二模态原始数据使用经典特征提取方法获取第二模态的低级表达T1,P1为n维向量,T1为1维向量;第一模态的低级表达P1通过堆叠的两层受限波尔兹曼机获得中级表达P3;第二模态的低级表达T1通过堆叠的两层受限波尔兹曼机获得中级表达T3;所述第一模态的中级表达P3与第二模态的中级表达T3分别使用自动编码器进行编码,编码后第一模态表示为P4,第二模态表示为T4,所述P4与T4为相同维数的向量,计算P4与T4的相似性测度。



1. 一种使用深层结构获取双模态相似性测度的方法,其特征在于,该方法包括:

第一模态原始数据使用经典特征提取方法获取第一模态的低级表达P1,第二模态原始数据使用经典特征提取方法获取第二模态的低级表达T1,其中,P1为n维向量,T1为l维向量;

第一模态的低级表达P1通过堆叠的两层受限波尔兹曼机获得中级表达P3,所述P3为s维二元向量;

第二模态的低级表达T1通过堆叠的两层受限波尔兹曼机获得中级表达T3,所述T3为s维二元向量;

所述第一模态的中级表达P3与第二模态的中级表达T3分别使用自动编码器进行编码,编码后第一模态表示为P4,第二模态表示为T4,所述P4与T4为相同维数的向量;

在编码层通过公式 $C(P3, T3; W_f, W_g) = ||f(P3; W_f) - g(T3; W_g)||_1$ 计算第一模态与第二模态的相似性测度C,其中 $|| \cdot ||_1$ 为L<sub>1</sub>范数, $f(P3; W_f) = P4$ , $g(T3; W_g) = T4$ , $W_f$ 为第一模态由P3计算得到P4所配置的所有偏置和连接权值的集合, $W_g$ 为第二模态由T3计算得到T4所配置的所有偏置和连接权值的集合。

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:

所述受限波尔兹曼机的配置参数包括可见层V神经单元 $v_i$ 的偏置 $b_i$ 、隐藏层H神经单元 $h_j$ 的偏置 $c_j$ 以及可见层神经单元和隐藏层神经单元的连接权值 $w_{ij}$ ;

所述配置参数为令可见层V和隐藏层H神经单元的联合概率分布 $p(v, h)$ 最大的数值,其中 $p(v, h) = \frac{1}{Z} \exp(-E(v, h))$ ,Z为归一化常数, $E(v, h)$ 是由受限波尔兹曼机的可见层神经单元、隐藏层神经单元的不同配置定义的能量函数。

3. 根据权利要求1所述的方法,其特征在于,所述 $W_f$ 与 $W_g$ 的训练算法进一步包括:

给定第一模态中级表达P3和第二模态中级表达T3,以及二元指标I,当训练用的第一模态与第二模态确实是同一对象的不同模态描述时,I取值为1,反之,第一模态与第二模态不是同一对象的不同模态描述时,I取值为0;

将权值集合 $W_f$ 与 $W_g$ 合并表示为 $\theta$ ,则定义任何成对输入的第一模态中级表达P3和第二模态中级表达T3的损失函数为:

$$L(P3, T3, I; \theta) = \alpha(L_I(P3; \theta) + L_T(T3; \theta)) + (1 - \alpha)L_C(P3, T3, I; \theta);$$

其中,

$$L_I(P3; \theta) = \|P3 - P5\|_2^2;$$

$$L_T(T3; \theta) = \|T3 - T5\|_2^2;$$

$$L_C(P3, T3, I; \theta) = IC^2 + (1 - I)\exp(-\lambda C);$$

$|| \cdot ||_2$ 为L<sub>2</sub>范数,P5为P3的重构表达,T5为对T3的重构表达, $L_I$ 和 $L_T$ 分别是两个子网络从编码到解码因数据的重构误差导致的损失;而 $L_C$ 是因第一模态和第二模态是否兼容而产生的对比损失;参数 $\lambda$ 是一个常数,由 $C(P3, T3; W_f, W_g)$ 在整个训练集上的上界确定;参数 $\alpha$ 为0到1之间的常数,用于平衡重构损失和兼容性损失。

## 一种使用深层结构获取双模态相似性测度的方法

### 技术领域

[0001] 本发明涉及多媒体检索技术,特别是一种使用深层结构获取双模态相似性测度的方法。

### 背景技术

[0002] 随着大规模计算与大数据的存储成为可能,多模态数据的学习与挖掘逐渐成为人们研究的热点。一方面,信息本身来源于多个模态,例如,网络上分享的旅行照片通常有些标记词汇,即照片包含了图像模态与文本模态的信息。另一方面,挖掘多源的异质数据有助于利用不同模态的信息,比如,说话人的关节和肌肉的运动通常有助于消除相似的发音引起的语音歧义。

[0003] 通常情况下,对于多模态数据的研究可以转化为对多个双模态数据的研究,这样从一定程度上可以降低直接研究多模态数据的复杂度,此前已有一些学习双模态数据的方法,其中包括,扩展潜在狄利克雷分派(LDA, Latent Dirichlet Allocation)模型,挖掘了图像和文本标签在主题级上的关系;建立图像和文本的联合模型,可以认为是LDA模型的无向图扩展;还有马尔科夫随机场(MRF, Markov Random Field)和LDA结合的模型。但是,以上三种模型只包含单个隐藏层,对于比较复杂的双模态数据则无法获得有效的表达。

[0004] 源于人脑的深层和分布式结构提出的深度学习试图学习获得层次化的、有效的表达,便于为后续的各种识别和挖掘提供更好的基础。在很短的发展中,深度学习在建模和挖掘单模态数据,如语音识别和计算机视觉等方面已经取得了瞩目的成就。

[0005] 受到深度学习的启发,一种使用深层的自编码器完成视觉与语音的融合任务的模型,提出了视频和语音两个模态数据的共同表示,即提取两种模态中的共同成分;还有一种生成模型的深层波尔兹曼机,提出学习图像和文本的共同表示。但是这两个模型都是期望获得不同模态数据的共同成分表示,主要集中于跨模态的检索,而不是计算他们之间的相似度;而基于双胞胎自编码器的框架,将多个模态融合到单一的表示空间,虽然可以完成多模态数据之间相似度的计算,但是这个框架只能用于有标记的数据,对于没有标记的数据则无法实现相似度的计算。

### 发明内容

[0006] 有鉴于此,本发明提供了一种使用深层结构获取双模态相似性测度的方法,应用深度学习框架,解决双模态数据的相似性测度计量问题。本发明提出的技术方案是:

[0007] 一种使用深层结构获取双模态相似性测度的方法,该方法包括:

[0008] 第一模态原始数据使用经典特征提取方法获取第一模态的低级表达P1,第二模态原始数据使用经典特征提取方法获取第二模态的低级表达T1,其中,P1为n维向量,T1为1维向量;

[0009] 第一模态的低级表达P1通过堆叠的两层受限波尔兹曼机获得中级表达P3,所述P3为s维二元向量;

[0010] 第二模态的低级表达T1通过堆叠的两层受限波尔兹曼机获得中级表达T3,所述T3为s维二元向量;

[0011] 所述第一模态的中级表达P3与第二模态的中级表达T3分别使用自动编码器进行编码,编码后第一模态表示为P4,第二模态表示为T4,所述P4与T4为相同维数的向量,计算P4与T4的相似性测度。

[0012] 上述方案中,所述相似性测度计算方法进一步包括:

[0013] 在编码层通过公式 $C(P3, T3; W_f, W_g) = ||f(P3; W_f) - g(T3; W_g)||_1$ 计算第一模态与第二模态的相似性测度C,其中 $|| \cdot ||_1$ 为L<sub>1</sub>范数, $f(P3; W_f) = P4$ , $g(T3; W_g) = T4$ , $W_f$ 为第一模态由P3计算得到P4所配置的所有偏置和连接权值的集合, $W_g$ 为第二模态由T3计算得到T4所配置的所有偏置和连接权值的集合。

[0014] 上述方案中,该方法进一步包括:

[0015] 所述受限波尔兹曼机的配置参数包括可见层V神经元 $v_i$ 的偏置 $b_i$ 、隐藏层H神经元 $h_j$ 的偏置 $c_j$ 以及可见层神经元和隐藏层神经元的连接权值 $w_{ij}$ ;

[0016] 所述配置参数为令可见层V和隐藏层H神经元的联合概率分布 $p(v, h)$ 最大的数值,其中 $p(v, h) = \frac{1}{Z} \exp(-E(v, h))$ ,Z为归一化常数,E(v, h)是由受限波尔兹曼机的可见层神经元、隐藏层神经元的不同配置定义的能量函数。

[0017] 上述方案中,所述 $W_f$ 与 $W_g$ 的训练算法进一步包括:

[0018] 给定第一模态中级表达P3和第二模态中级表达T3,以及二元指标I,当训练用的第一模态与第二模态确实是同一对象的不同模态描述时,I取值为1,反之,第一模态与第二模态不是同一对象的不同模态描述时,I取值为0;

[0019] 将权值集合 $W_f$ 与 $W_g$ 合并表示为 $\theta$ ,则定义任何成对输入的第一模态中级表达P3和第二模态中级表达T3的损失函数为:

[0020]  $L(P3, T3, I; \theta) = \alpha(L_I(P3; \theta) + L_T(T3; \theta)) + (1 - \alpha)L_C(P3, T3, I; \theta)$ ;

[0021] 其中,

[0022]  $L_I(P3; \theta) = ||P3 - P5||_2^2$ ;

[0023]  $L_T(T3; \theta) = ||T3 - T5||_2^2$ ;

[0024]  $L_C(P3, T3, I; \theta) = IC^2 + (1 - I)\exp(-\lambda C)$ ;

[0025]  $|| \cdot ||_2$ 为L<sub>2</sub>范数,P5为P3的重构表达,T5为对T3的重构表达, $L_I$ 和 $L_T$ 分别是两个子网络从编码到解码因数据的重构误差导致的损失;而 $L_C$ 是因第一模态和第二模态是否兼容而产生的对比损失;参数 $\lambda$ 是一个常数,由 $C(P3, T3; W_f, W_g)$ 在整个训练集上的上界确定;参数 $\alpha$ 为0到1之间的常数,用于平衡重构损失和兼容性损失。

[0026] 综上所述,本发明技术方案提出了一种可计算的深层结构,用于度量双模态数据的相似性,本发明提出的框架将单个模态中特征提取方法与堆叠受限波尔兹曼机模型的深层神经网络结构结合,通过自动编码器学习双模态数据之间的相似度。

## 附图说明

[0027] 图1为本发明进行双模态相似性测度测量的深度学习框架。

[0028] 图2为受限波尔兹曼机神经网络结构图。

### 具体实施方式

[0029] 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。

[0030] 本发明一个实施例的技术方案是:

[0031] 第一模态原始数据使用经典特征提取方法获取第一模态的低级表达P1,第二模态原始数据使用经典特征提取方法获取第二模态的低级表达T1,其中,P1为n维向量,T1为1维向量;

[0032] 第一模态的低级表达P1通过堆叠的两层受限波尔兹曼机获得中级表达P3,所述P3为s维二元向量;

[0033] 第二模态的低级表达T1通过堆叠的两层受限波尔兹曼机获得中级表达T3,所述T3为s维二元向量;

[0034] 所述第一模态的中级表达P3与第二模态的中级表达T3分别使用自动编码器进行编码,编码后第一模态表示为P4,第二模态表示为T4,所述P4与T4为相同维数的向量,计算P4与T4的相似性测度。

[0035] 考虑深度学习的进展,本发明提出了一种使用深层结构获取双模态相似性测度的方法。本发明提出的双模态深度结构称为BDA(Bimodal Deep Architecture),它包含三个密切关联的组件,如图1所示。下面详细描述三个组件的细节。

[0036] 第一组件:

[0037] 第一组件可以使用各自单个模态中经典的特征提取方法来获取第一模态的低级表达P1与第二模态的低级表达T1,单模态的经典特征提取方法为现有技术,且发展较为成熟,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。第一组件得到的所述低级表达P1为n维向量,T1为1维向量,通常n不等于1,且P1与T1的元素数据类型也可能不同,即一般情况P1与T1不能直接用于相似度计算。

[0038] 第二组件:

[0039] 第二组件为堆叠受限波尔兹曼机(RBM,Restricted Boltzmann Machine)的神经网络模型,如图1中的第二组件所示,第一模态与第二模态均堆叠了两层受限波尔兹曼机,每层受限波尔兹曼机的神经网络结构都相同,只是配置参数有所不同,图2为受限波尔兹曼机的神经网络结构图:可见层V包含m个神经元 $v_1 \sim v_m$ ,每个神经元 $v_i$ 的偏置为 $b_i$ ,可见层神经元之间没有连接;隐藏层H包含s个神经元 $h_1 \sim h_s$ ,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2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。

[0040] 对于第一模态的第一层RBM,其可见层神经元的个数m等于第一模态低级表达P1的维数n,即 $m=n$ ,所述第一模态低级表达P1的n个向量元素 $p_{11} \sim p_{1n}$ 与第一模态的第一层RBM可见层 $v_1 \sim v_m$ 的取值一一对应;同样的,对于第二模态的第一层RBM,其可见层神经元的个数m等于第二模态低级表达T1的维数1,即 $m=1$ ,所述第二模态低级表达T1的1个向量元素 $t_{11} \sim t_{11}$ 与第二模态的第一层RBM可见层 $v_1 \sim v_m$ 的取值一一对应。

[0041] 对于第一模态的第二层RBM,其隐藏层神经元 $h_1 \sim h_s$ 输出值与图1中第一模态的

中级表达P3的向量元素 $p_{31} \sim p_{3s}$ 一一对应;同样的,对于第二模态的第二层RBM,其隐藏层神经单元 $h_1 \sim h_s$ 输出值与图1中第二模态的中级表达T3的向量元素 $t_{31} \sim t_{3s}$ 一一对应,P3与T3向量维数相同。

[0042] 每层受限波尔兹曼机的可见层神经单元的偏置 $b_i$ 、隐藏层神经单元的偏置 $c_j$ 以及可见层神经单元和隐藏层神经单元的连接权值 $w_{ij}$ 为训练算法计算得到的配置参数,经过大量双模态数据对本发明提出的深度学习框架反复训练得到的,最终训练得到的所述配置参数能够使得整个深度学习框架具有稳定良好的性能。

[0043] 下面介绍所述受限波尔兹曼机的可见层神经单元的偏置 $b_i$ 、隐藏层神经单元的偏置 $c_j$ 以及可见层神经单元和隐藏层神经单元的连接权值 $w_{ij}$ 的训练算法:

[0044] 如图2所示,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:

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

[0046] 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,例如,

[0047] 当可见层神经单元和隐藏层神经单元都依伯努利分布时:

$$E(v, h) = -\sum_{i=1}^m \sum_{j=1}^s w_{ij} v_i h_j - \sum_{i=1}^m b_i v_i - \sum_{j=1}^s c_j h_j \quad (2)$$

[0049] 当可见层神经单元依高斯分布、隐藏层神经单元依伯努利分布时:

$$E(v, h) = -\sum_{i=1}^m \sum_{j=1}^s w_{ij} \frac{v_i}{\sigma_i} h_j + \sum_{i=1}^m \frac{(v_i - b_i)^2}{2\sigma_i^2} - \sum_{j=1}^s c_j h_j \quad (3)$$

[0051] 其中所有可见层神经单元的方差 $\sigma_i$ 设置为1。

[0052] 当第二模态为文本模态时,在第一组件中用词袋模型对其进行特征提取,得到文本模态的低级表达P1,此时将第二组件的第一层受限波尔兹曼机可见层神经单元设计为依重复的Softmax分布、隐藏层神经单元依伯努利分布时,则相应的能量函数为:

$$E(v, h) = -\sum_{i=1}^m \sum_{j=1}^s w_{ij} v_i h_j - \sum_{i=1}^m b_i v_i - M \sum_{j=1}^s c_j h_j \quad (4)$$

[0054] 其中,P1的编码原则为,当文本模态原始数据包括词典中的高频词时编码为1,否则为0,M为P1非零元素的个数。

[0055] 以上给出了几个可见层和隐藏层的能量函数定义,但不限于以上几种能量函数定义,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。

[0056] 如果RBM的可见层神经单元的偏置 $b_i$ 、隐藏层神经单元的偏置 $c_j$ 以及可见层神经单元和隐藏层神经单元的连接权值 $w_{ij}$ 的取值能够使可见层V和隐藏层H神经单元的联合概率分布 $p(v, h)$ 最大,也就是 $E(v, h)$ 最小,则认为所述 $b_i$ 、 $c_j$ 、 $w_{ij}$ 的取值使得本发明深度学习框架的性能最好。

[0057] 进一步的,为了便于计算,本发明将第一模态的第二层受限波尔兹曼机与第二模态的第二层受限波尔兹曼机设计为伯努利-伯努利型受限波尔兹曼机,即上述两个第二层

RBM的可见层神经单元依伯努利分布,隐藏层神经单元依伯努利分布;由于第一模态的第一层受限波尔兹曼机的隐藏层即为第一模态的第二层受限波尔兹曼机的可见层,因此将第一模态的第一层受限波尔兹曼机设计为X-伯努利型受限波尔兹曼机,其中X代表任意概率模型,根据第一模态低级表达P1确定,例如当P1的元素为连续型实值时,X为高斯型,当P1的元素为0、1时,X为重复的Softmax,第一模态第一层RBM的隐藏层神经单元依伯努利分布;同理,第二模态的第一层RBM也设计为X-伯努利型受限波尔兹曼机,原因与第一模态第一层RBM的设计相同,在此不再赘述。

[0058] 综上可知,第一组件得到的第一模态低级表达P1与第二模态低级表达T1经过第二组件的堆叠受限波尔兹曼机,可以分别得到第一模态的中级表达P3与第二模态的中级表达T3,所述P3与T3的维数相同,且均为二元表示,即0、1表示。

[0059] 第三组件:

[0060] 本申请深度学习框架中的第三个组件提出了自动编码器用于双模态的表示,以学习双模态数据之间的相似度,如图1所示。

[0061] 第三组件中提出的自动编码器包含两个子网络,第一子网络对第一模态中级表达P3进行编码得到P4,再对P4进行解码得到P5,其中P5为第一模态中级表达P3的重构表达,即近似表达,与P3存在一定的重构误差;第二子网络对第二模态中级表达T3进行编码得到T4,再对T4进行解码得到T5,其中T5为第二模态中级表达T3的重构表达,即近似表达,与T3存在一定的重构误差。每个子网络都是全连接的感知器。而在编码层,两个子网络通过预先定义的相似性测度连接。通过引用基于能量学习的方法设计适当的损失函数,能够学习到两个模态数据的相似性。

[0062] 从输入到编码层的两个子网络分别记为 $f(P3;W_f)$ 和 $g(T3;W_g)$ ,其中, $f(\cdot)$ 表示第三组件中第一模态中级表达P3到P4的变换法则,即 $P4=f(P3;W_f)$ ,其中第三组件第一模态从输入到编码层的网络结构与第二组件的BB-RBM神经网络结构相同,P3为可见层,P4为隐藏层,分别对应可见层神经单元偏置、隐藏层神经单元偏置以及可见层神经单元和隐藏层神经单元的连接权值,将上述第三组件从输入到编码层的所有偏置和连接权值用 $W_f$ 表示;同样的, $g(\cdot)$ 表示第三组件中第二模态中级表达T3到T4的变换法则,即 $T4=g(T3;W_g)$ ,且 $W_g$ 的定义与 $W_f$ 相同,在此不再赘述。P4与T4的维数相同,且为二元表示。

[0063] 在编码层对P4与T4进行相似性测度计算:

[0064]  $C(P4,T4)=||P4-T4||_1$  (5)

[0065] 又因为 $P4=f(P3;W_f)$ , $T4=g(T3;W_g)$ ,相似性测度公式又可以表示为:

[0066]  $C(P3,T3;W_f,W_g)=||f(P3;W_f)-g(T3;W_g)||_1$  (6)

[0067] 其中 $||\cdot||_1$ 为 $L_1$ 范数。

[0068] 由此计算出的相似性测度C即为两个模态的相似程度,以此用来描述这两个模态描述同一个对象时的匹配程度。

[0069] 上述提到的 $W_f$ 与 $W_g$ 是本申请第二组件经过训练算法计算得到的配置参数,是在大量双模态数据反复训练后得出的,是在双模态数据进行相似性测度计算之前预先配置好的参数。下面介绍 $W_f$ 与 $W_g$ 的训练算法。

[0070] 在 $W_f$ 与 $W_g$ 的训练算法中,本发明提出了一个损失函数,给定输入对第一模态P3和第二模态T3,以及二元指标I,二元指标I表示第一模态和第二模态是否匹配,当训练用的第一

模态与第二模态确实是同一对象的不同模态描述时,  $I$  取值为1; 反之, 第一模态与第二模态不是同一对象的不同模态描述时,  $I$  取值为0。为简化记号, 将配置参数集合  $W_f$  与  $W_g$  合并表示为  $\theta$ 。则可以定义任何成对输入的第一模态和第二模态的损失函数为:

$$[0071] \quad L(P3, T3, I; \theta) = \alpha(L_I(P3; \theta) + L_T(T3; \theta)) + (1 - \alpha)L_C(P3, T3, I; \theta) \quad (7)$$

[0072] 其中,

$$[0073] \quad L_I(P3; \theta) = \|P3 - P5\|_2^2 \quad (8)$$

$$[0074] \quad L_T(T3; \theta) = \|T3 - T5\|_2^2 \quad (9)$$

$$[0075] \quad L_C(P3, T3, I; \theta) = IC^2 + (1 - I)\exp(-\lambda C) \quad (10)$$

[0076] 这里,  $\|\cdot\|_2$  为  $L_2$  范数,  $P5$  为  $P3$  的重构表达,  $T5$  为  $T3$  的重构表达,  $L_I$  和  $L_T$  分别是两个子网络从编码到解码因数据的重构误差导致的损失。而  $L_C$  是因第一模态和第二模态是否兼容而产生的对比损失。方程(10)中的参数  $\lambda$  是一个常数, 由  $C(P3, T3; W_f, W_g)$  在整个训练集上的上界确定。方程(7)中的参数  $\alpha$  为0到1之间的常数, 用于平衡重构损失和兼容性损失。

[0077]  $W_f$  与  $W_g$  为令公式(7)中的损失函数最小的配置。

[0078] 整个自动编码器的学习可以使用标准的后传算法得到。在学习过程中, 两个子网在编码层通过相似性测度耦合。因此, 对于新输入的第一模态与第二模态, 通过学习到的权值参数, 这个自动编码器可以得到对应的编码  $T4$  与  $P4$ 。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。



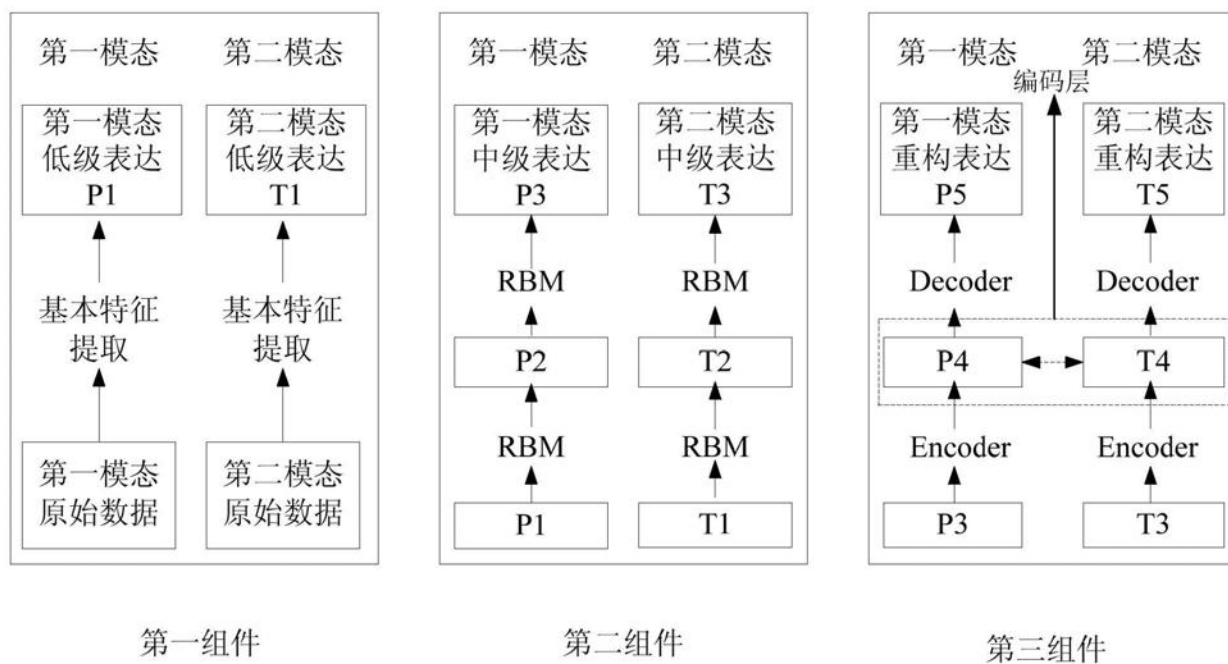


图1

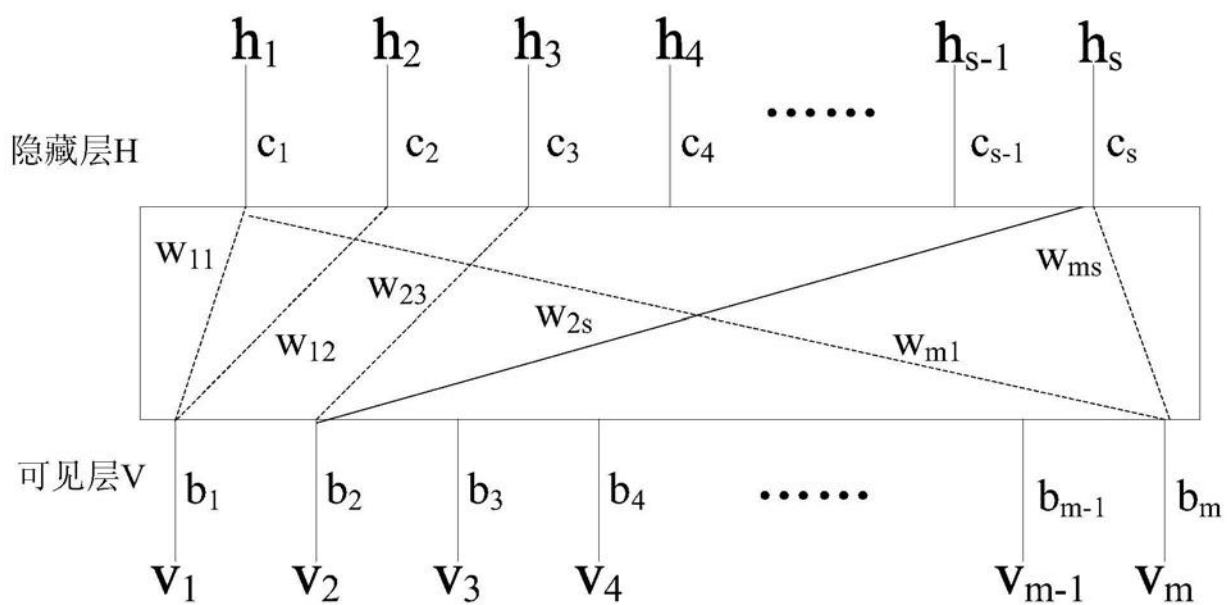


图2