本周主要工作：学习基于受限玻尔兹曼机的跨模态信息检索

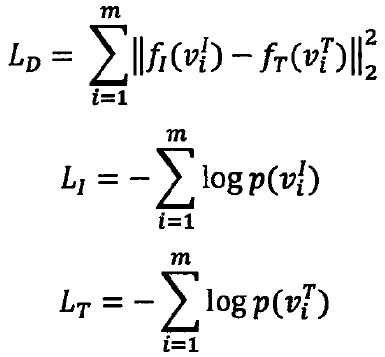
作者开始提到了自编码器和受限玻尔兹曼机两种非常重要的模型，前者是确定性模型，后者是利率模型。由于在多个层次上建立不同模态数据的关联在自编码器上还没有比较好的解决方案，而且受限玻尔兹曼机提供了这种可能。

作者首先提出对应首先玻尔兹曼机——Corr-RBM，然后又提出了对应深度信念网络——Corr-DBN和栈式对应受限玻尔兹曼机——Corr-RBMs。

Corr-RBM



这个模型包含了两个单模态的RBM，神经元目相同，而且只在最顶层即表示层进行关联约束。设fI和fT分别是两个模态的从输入层到输出层的映射函数，那么设参数，那么要训练数据的目标是满足目标函数公式的参数θ。其中目标函数的具体内容是

（v为输入数据，α、β是图像文本似然在目标函数的比重），目的是让关联误差和似然和最小。最小化关联误差会使得图像和文本在表示空间中更近，而最小化似然会图像和文本数据再表示空间中保持区分性。

参数训练算法（先单独用对比散度算法来更新两个模态的参数，然后用梯度下降法来依据关联误差来更新参数）

Corr-RBM学习算法

1. 用对比散度方法来更新图像RBM参数（CD-K）

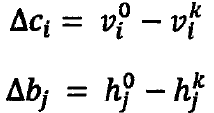


1. 用对比散度法来更新文本RBM参数（CD-K）

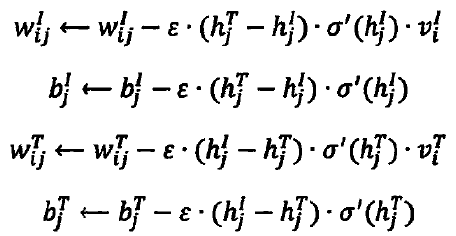


其中ɛ是学习速率，θ的具体计算方法如下：





1. 用梯度下降法根据关联误差来更新参数



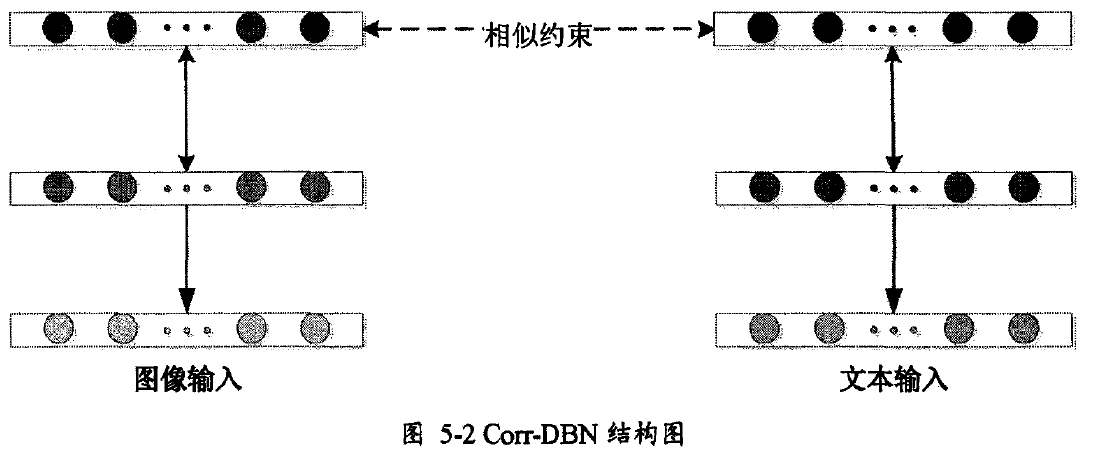
其中h是表示层单元，б是激活函数

1. 重复步骤1、2、3直至收敛

更加深层的结构

对应深度信念网Corr-DBN

通过栈式叠加RBM和Corr-RBM得到。

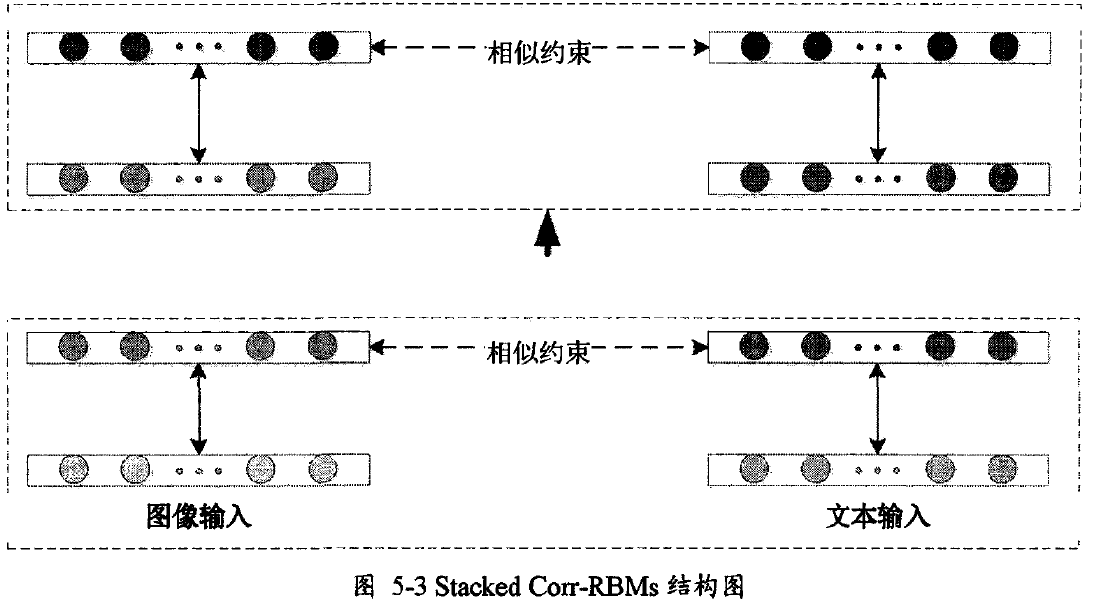


那么先通过叠加多个单模态的图像和文本RBM学习更高层次的单模态数据表示，然后在最高层使用Corr-RBM来学习图像和文本的跨模态关联。训练算法如下：先用CD-K训练除去最顶层的单模态RBM，然后用Corr-RBM学习算法来训练顶层的Corr-RBM。

栈式对应受限玻尔兹曼机（Stacked Corr-RBMs）

前面提到的Corr-RBM和Corr-DBN都只是在最高层学习不同模态的关联信息，称之为“一次关联模型”故而提出在多层之间添加相似约束的“多次关联模型”。

通过叠加多个Corr-RBM来得到。



训练算法（分多层次）：

设（VI，VT）是最底层的图像和文本输入表示，N个表示表示层（hIj，hjT）（2<=j<=N分别自底向上表示表示层图像和文本。

首先用算法Corr-RBM学习算法训练由输入（VI，VT）和表示层（hI1，hT1）上组成的Corr-RBM，然后继续用这个算法训练由（hI2，hT2）（hI2，hT2）组成的Corr-RBM，直至所有层训练完毕。

其他学习：

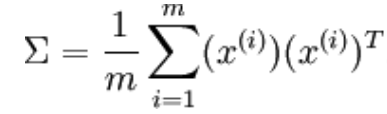
PCA主成分分析

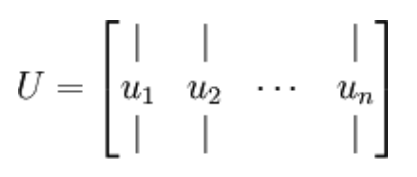
PCA是一种能极大化提升无监督特征学习速度的数据将为算法，而且有益于实现白化算法。

准备阶段：数据预处理，让数据具有相同的均值（0）和方差。

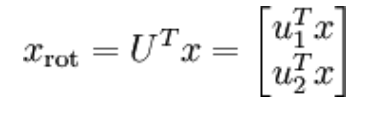
PCA算法寻找一个低维的空间来投影我们的数据

如果均值为零的话，那么

Σ称为协方差矩阵，那么数据变化的主方向就是协方差矩阵 的主特征向量（选取特征值最大的λ对应的特征向量u），那么把所有特征向量按列来摆放形成一个矩阵U

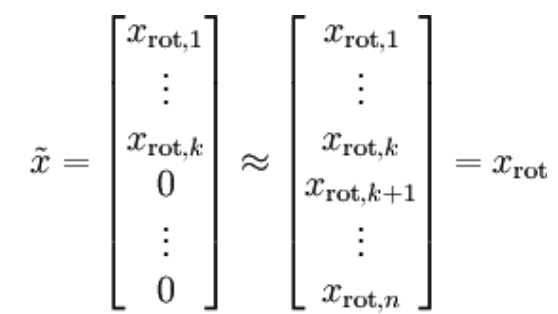


旋转数据：



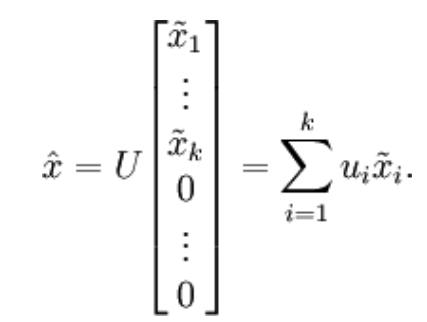
这样就得到了数据在特征向量方向上的投影，若要还原数据

假如想把数据降到K维表示,只需选取Xrot的前K个成分，分别对应前K个数据变化的主方向。实际上，PCA算法做的其实就是丢弃Xrot中后面（取值较小）的成分，就是将这些成分的值近似为零。具体的说，设X（矩阵）是Xrot的近似表示，那么那么将Xrot除了前K个成分外，其余全赋值为零，就得到：



还原近似数据：

我们可以这样计算：

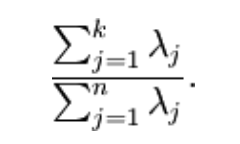


但是实际上我们不这样做，因为会有许多我用的乘0操作，而是计算前K维，后面补上N-K个零。

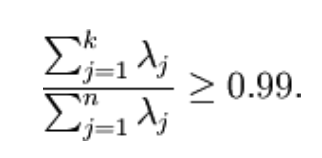
选择主成分的个数

定义保留的方差百分比如下：





经验的保留方差百分比为99%，选取的K值满足



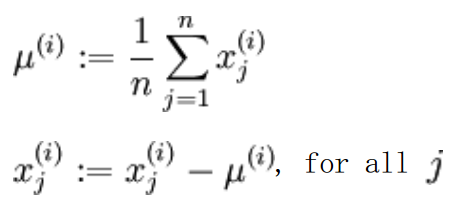
对图像数据应用PCA算法

对于自然图像：

为使PCA算法正常工作，我们通常需要满足以下要求：(1)特征的均值大致为0；(2)不同特征的方差值彼此相似。对于自然图片，即使不进行方差归一化操作，条件(2)也自然满足，故而我们不再进行任何方差归一化操作（对音频数据,如声谱,或文本数据,如词袋向量，我们通常也不进行方差归一化）。实际上，PCA算法对输入数据具有缩放不变性，无论输入数据的值被如何放大（或缩小），返回的特征向量都不改变。更正式的说：如果将每个特征向量X都乘以某个正数（即所有特征量被放大或缩小相同的倍数），PCA的输出特征向量都将不会发生变化。 唯一还需进行的规整化操作就是均值规整化，其目的是保证所有特征的均值都在0附近。根据应用，在大多数情况下，我们并不关注所输入图像的整体明亮程度。比如在对象识别任务中，图像的整体明亮程度并不会影响图像中存在的是什么物体。更为正式地说，我们对图像块的平均亮度值不感兴趣，所以可以减去这个值来进行均值规整化。

均值化处理：

如果X代表图像的亮度（灰度），那么就可以用这个算法进行规整化处理



请注意：1）对每个输入图像块Xi都要单独执行上面两个步骤，2）这里ui是指图像块X的平均亮度值。尤其需要注意的是，这和为每个像素Xj单独估算均值是两个完全不同的概念