**第一篇(基于流形特征的视频情感分析与识别)**

**资料文献：**

李青.蒋冬梅.FAN Ping.Ravyse Ilse.Sahli Hichem.LI Qing.JIANG Dong-mei.FAN Ping.Ravyse Ilse.Sahli Hichem 基于流形特征的视频情感分析与识别[期刊论文]-计算机工程与科学 2010(12)

**核心问题：**

目前视觉情感的提取一般关注于静态情感特征提取，但是对情感的变化无法进行细节的描述与分析。如何针对连续情感描述，捕捉更加丰富的情感信息是本文的核心关注问题。对于连续的帧数变换过程就会涉及到降维，降维当然意味着信息的丢失，不过鉴于实际数据本身常常存在的相关性，我们可以想办法在降维的同时将信息的损失尽量降低。

**核心思想&解决方案：**

本文提出将由各种情感特征串接起来的高维向量进行一种特别的特征降维，并将降为后的特征空间与情感描述空间关联起来，使得将为特征能够描述情感强度的细致变化。

* 原始视觉情感特征提取
* 视觉情感降维

将提取的描述心情的人脸七个器官动作作为情感分析的七个维度，；将以情歌维度降维到三维。本文提到的Isomap算法，将多维的信息点降维到三维，以在平面中可以显示出来信息点的拟合曲线走向，进而和六种基本情感建立一种映射关系。

* 视觉情感流形特征和情感分析
* 视觉情感识别

**价值&缺陷：**

本文提出的Isomap算法对高维情感数据值进行降维，从实验效果来看，比PCA降维算法的降维结果更加平滑，能将连续帧之间的情感变化更好的表示出来。

但是，在视觉情感识别的第一步，本文提出根据初识帧作为基准帧，随后的帧数特征值都是和这个基准帧做差得到的数据；那么这个初始帧的情感如何保证是中性的，还需要仔细讨论。

**注解：**

1）MDS：多维尺度化，

多维尺度分析Multidimensional Scaling，简称MDS，是一种探索性数据分析技术，主要是用适当的降维方法，将多个变量通过坐标定位在低维空间中(二维或三维)，变量之间的欧氏距离就可以反映它们之间的差异性和相似性。

2）PCA降维方法：

Principal Component Analysis(PCA)是最常用的线性降维方法，它的目标是通过某种线性投影，将高维的数据映射到低维的空间中表示，并期望在所投影的维度上数据的方差最大，以此使用较少的数据维度，同时保留住较多的原数据点的特性。

**第二篇(基于lucenen的网络视频垂直搜索系统设计与实现)**

**资料文献：**

**左军. 基于Lucene网络视频垂直搜索系统的设计与实现[D]. 北京邮电大学, 2007.**

**核心问题：**

针对全文搜索，传统的sql语句like性能太低。为了高效搜索文本内容大，文本数量大的问题，需要给出一套高效的，分词+索引的搜索系统。

**核心思想&解决方案：**

本文提出了一套解决搜索问题的框架，针对文本量大的问题，首先对其进行文本分词；随后根据分词的结果建立数据库索引，使得基于索引的搜索效率很大的提高。

**价值&缺陷：**

**lucenen框架的主要价值是提高检索的效率。**

**首先，分词机制采用了词表切分取代自动切分，利用完善的词表，大大降低了分词之后的单词文件大小。**

**其次，检索机制，传统的数据库需要逐个遍历记录机型模糊匹配，必有索引的搜索速度要有多个数量级的下降。所以采用索引的搜索引擎很有必要。并且基于lucenen的搜索建立在分词的基础上，会根据匹配度的大小将数据依次返回回来，搜索结果范围显著提高。**

**但是事实检验，如果针对文本量不是很大的搜索对象时，搜索的效率提升与数据库索引搜索相比，提升的幅度有限。**

**注解：**

反向索引：反向索引就是说我们维护了一个词 / 短语表，对于这个表中的每个词 / 短语，都有一个链表描述了有哪些文档包含了这个词 / 短语。这样在用户输入查询条件的时候，就能非常快的得到搜索结果。

**第三篇(多线程爬虫)**

**资料文献：**

**核心问题：**

**核心思想&解决方案：**

**价值&缺陷：**

**注解：**

**第 篇()**

**资料文献：**

**核心问题：**

**核心思想&解决方案：**

**价值&缺陷：**

**注解：**

**第 篇()**

**资料文献：**

**核心问题：**

**核心思想&解决方案：**

**价值&缺陷：**

**注解：**