
隐马尔可夫模型实现语音和视频识别*

林文永, 石志国, 薛为民, 陈锋军

(北京科技大学信息工程学院, 北京 100083)

E-mail: tre2@sohu.com

摘要: 利用隐马尔可夫模型(HMM)对多媒体数据仓库进行复杂数据挖掘,复杂数据挖掘要解决的难题就是音频和视频识别。建立音频和视频的识别模型及其相关的算法,在视频识别算法上,构造出符合HMM的识别方法。根据模型建立系统,实验证明声音的识别率最高达到96.67%,视频中特征值的检测率可以达到87.81%。研究结果可以应用在多媒体的识别和数据挖掘领域,提供一个比较完整的复杂数据挖掘的模型和算法。

关键词: HMM; 复杂数据挖掘; 音频识别; 视频检测关

作者简介: 林文永(1979-),男,浙江温州人,硕士研究生,主要研究方向为人工智能;石志国(1978-),男,江苏泗阳人,硕士研究生,主要研究方向为无线通信、图像语音信号处理;薛为民(1968-),男,博士研究生,主要研究方向为语音图像识别与跟踪;陈锋军(1977-),女,硕士,研究方向为多媒体处理技术。

Hidden Markov Model Based the Recognition of Audio and Video

Lin Wen-yong, Shi Zhi-guo, Xue Wei-min, Chen Feng-jun

(Information Engineering School, University of Science & Technology Beijing, Beijing 100083 China)

E-mail: tre2@sohu.com

Abstract: In this paper, it mainly discusses a method to process complex data mining in the multimedia data ware by HMM (Hidden Markov Model). The problem of complex data mining is the recognition of the audio and video. This paper will build the model of recognition of audio and video. It will construct the method according to the HMM. Due to results of experiments, the highest ratio of recognition of audio is 96.67%, the highest ratio of recognition of video is 87.81%. The application area of the results covers recognition of multimedia and the fields of data mining. This paper will provide completely models and technique of CDM (Complex Data Mining).

Key words: HMM; Complex Data Mining; Audio Recognition; Video Recognition

* 基金项目:“十五”国家科技攻关项目(2001BA605A)