

<<时空视频检索>>

图书基本信息

书名：<<时空视频检索>>

13位ISBN编号：9787811336115

10位ISBN编号：7811336111

出版时间：2010-5

出版时间：哈尔滨工程大学出版社

作者：任伟

页数：212

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

## 前言

The problem of semantic video scene categorisation by using spatio-temporal information is one of the significant open challenges in the field of video retrieval. During the past few years, advances in digital storage technology and computer performance have promoted video as a valuable information resource. Numerous video retrieval techniques have been successfully developed. Most of the techniques for video indexing and retrieval have extended the previous work in the context of image based retrieval. In this process, video sequences are treated as collections of still images. Relevant key-frames are first extracted followed by their indexing using existing image processing techniques based on low-level features. For the research in the book the key question is how to encode the spatial and temporal information in video for its efficient retrieval. Novel algorithms are proposed for matching videos and are compared them with state-of-the-art. These algorithms take into account image objects and their spatial relationships, and temporal information within a video which correlates with its semantic class. Also, the algorithms perform hierarchical matching starting with frame, and shot level before overall video level similarity can be computed. The approach, then, is exhaustively tested on the basis of precision and recall measures on a large number of queries and use the area under the average precision recall curve to compare the methods with those in the literature. As a part of this book an international video benchmark Minerva was proposed on which the results have been discussed.

## <<时空视频检索>>

### 内容概要

本书重点挖掘了视频的时空关系，探索了利用机器学习的方法进行视频切割、语义分类。

本书分七章，阐明了图像的各种特性，论述了视频的特征，系统介绍了视频的时空逻辑关系、视频的统计分析方法，研究了如何捕捉视频的时空特性，如何利用人工智能神经网络进行视频切割，如何训练计算机“学会”用人类的思维进行视频语义分类、检索。

各章节撰写排列体现了从简到繁、由浅入深、从理论到实际、从技术到系统的特点。

本书可以作为高等学校信号与图像处理、计算机科学、机器学习、人工智能、机器视觉等领域的研究生教材和参考书，也可以作为在这些领域从事相关工作的高级科学技术人员的参考书。

## 书籍目录

Chapter I Introduction 1.1 Motivation 1.2 Proposed Solution 1.3 Structure of BookChapter II Approaches to Video Retrieval 2.1 Introduction 2.2 Video Structure and Properties 2.3 Query 2.4 Similarity Metrics 2.5 Performance Evaluation Metrics 2.6 SystemsChapter III Spatio-temporal Image and Video Analysis 3.1 Spatio-temporal Information for Video Retrieval 3.2 Spatial Information Modelling in Multimedia Retrieval 3.3 Temporal Model 3.4 Spatio-temporal Information FusionChapter IV Video Spatio-temporal Analysis and Retrieval (VSTAR). A New Model 4.1 VSTAR Model Components 4.2 Spatial Image Analysis 4.3 A Model for the Temporal Analysis of Image Sequences 4.4 Video Representation, Indexing, and Retrieval Using VSTAR 4.5 Conclusions Chapter V Two Comparison Baseline Models for Video Retrieval 5.1 Baseline Models 5.2 Adjero et al. (1999) Sequences Matching--Video Retrieval Model 5.3 Kim and Park (2002a) data set matchingmVideo Retrieval Model Chapter VI Spatio-temporal Video RetrievalmExperiments and Results 6.1 Purpose of Experiments 6.2 Data Description 6.3 Spatial and Temporal Feature Extraction 6.4 Video Retrieval Models: Procedure for Parameter Optimisation 6.5 Video Retrieval Models:Results on Parameter Optimisation 6.6 Comparison of Four Models 6.7 Model Robustness (Noise) 6.8 Computational Complexity 6.9 ConclusionsChapter VII Conclusions 7.1 Reflections on the book as a whole.....Reference

## 章节摘录

插图：In Ioka and Kurokawa ( 1992 ) , the user is allowed to specify a query by drawing a motion trajectory. The similarity is computed as the Euclidean distance between the query vector and the stored vector for each given interval to match the specified trajectory with the trajectories of the sequences in the database.

3.3.2.2 Correlation Based Comparison This approach is based on finding the maximum correlation between the predictor and the current one, for gesture recognition to identify actions. Martin and Shah ( 1992 ) used dense optical flow fields over a region, and computed correlation between different sequences for matching. In Campbell and Bobick's ( 1995 ) work on gesture recognition, the learning/training process is accomplished by fitting the unique curve of a gesture into the subset of the phase space with low-order polynomials. Rui and Anandan ( 2000 ) addressed the problem of detecting action boundaries in a video sequence containing unfamiliar and arbitrary visual actions. Their approach was based on detecting temporal discontinuities of the spatial pattern of object region motion which correspond to the action temporal boundary to capture the action. They represented frame-to-frame optical flow in terms of the coefficients calculated from all of the flow fields in a sequence, after principal components analysis to determine the most significant such flow fields. The temporal trajectories of those coefficients of the flow field are analysed to determine locations of the action segment boundaries of video objects.

## <<时空视频检索>>

### 编辑推荐

《时空视频检索(英文版)》：学者书屋系列

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>