

分类号_____

密级_____

UDC _____

编号_____

中国科学院研究生院

硕士学位论文

基于时空信息的关注分析技术研究及其应用

邱学侃

指导教师_____黄庆明_____教授_____

中国科学院研究生院信息科学与工程学院

申请学位级别_____工学硕士_____ 学科专业名称_____计算机应用技术_____

论文提交日期_____2009年4月_____ 论文答辩日期_____2009年5月_____

培养单位_____中国科学院研究生院（本部）_____

学位授予单位_____中国科学院研究生院_____

答辩委员会主席_____

声 明

我声明本论文是我本人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，本论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

作者签名：

日期：

论文版权使用授权书

本人授权中国科学院研究生院可以保留并向国家有关部门或机构送交本论文的复印件和电子文档，允许本论文被查阅和借阅，可以将本论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编本论文。

（保密论文在解密后适用本授权书。）

作者签名：

导师签名：

日期：

摘 要

随着计算机技术的发展，网络传输带宽不断提高，存储成本日益下降，越来越多的数码便携设备也进入到人们的日常生活中，这些都促使了数字媒体信息的飞速膨胀。因此，如何处理这些海量的图像和视频数据已经成为现在的研究热点。关注分析作为一个新的研究方向，其目的是基于人类的视觉感知特性，使计算机能够自动获得图像和视频关注的时空分布。利用关注分布，一系列的应用可以相继展开，比如视频摘要提取，小屏幕设备的视频浏览，图像和视频的虚拟内容插入，以及可变码率的视频压缩等等。

本文针对关注分析这个方向展开研究，提出了一个基于空域信息的图像关注分析方法和一个融合多种时空特征的视频关注分析方法，并在此基础上进行了视频时空摘要提取的应用。

第一，本文提出了一个基于空域信息的图像关注分析方法，该方法以区域作为图像关注分析的感知单元，并使用综合了多种空域特征，包括空间对比度、信息量、区域的位置和面积等，使其更加符合人类的视觉感知特性。该方法最终可以得到图像关注的分布情况，即标识各个像素和区域关注度的关注度图。

第二，本文提出了一个融合了多种时空特征的视频时空关注分析方法，该方法的空域模型使用了图像关注分析方法，同时综合了视频的多种时域特征，包括摄像机运动、局部运动、时域对比度等，通过分析得到视频的时空关注分布情况，即空域上的关注度图以及时域上标识各帧关注度的关注变化曲线。

第三，基于该视频关注分析方法，本文提出了一个视频的时空摘要提取方法。所谓的时空摘要，即从时域和空域两方面来提取视频的摘要部分，以满足手机等小屏幕便携设备的视频浏览需求。以关注分析为基础，该方法首先提取关注较高的片段作为时域摘要，然后在时域摘要的基础上进一步提取空域的关注区域，并作为最终的时空摘要。与传统的摘要技术不同的是，该方法能够从时空两方面来提取视频摘要，同时，该方法的基础是视频的时空关注分布，即能够以符合人类视觉感知的方式提取关注度较高的时空子集作为摘要。

关键词：关注分析 视频时空摘要提取

Research and Application of Attention Analysis Techniques Based on Spatio-temporal Information

QIU XUEKAN(Computer Application)

Supervised by: **HUANG Qing-ming**(Professor)

With the growth of computer techniques, the speeding up of the network bandwidth, the decreasing of storage cost, and the increasing of digital devices in the daily life, more and more digital media data is emerging into the world. So, how to manage the giant multimedia information has become a hot research point. Attention analysis, which is a new research domain and is based on human visual perception, is to automatically obtain the attention distribution of image and video by computer. With the attention distribution of images and videos, many applications would be carried out, such as video summarization, video browsing on small screens, virtual content insertion into images and videos, video compression, and so on.

This paper describes our research work on attention analysis. Our work is mainly carried out on three aspects, an image attention analysis method, a video attention analysis method which combines multiple spatial and temporal features, and a spatio-temporal video summarization method based on attention analysis.

Firstly, in this paper, we propose an image attention analysis method which is based on spatial information. This method takes region as perceptive unit, and combines multiple features including spatial contrast, information, location and area of region, and so on. This method accords with human visual perception. Using this method we can obtain the saliency map, which is the spatial distribution of image attention, denoting the attention value of each pixel.

Secondly, we propose a spatio-temporal video attention analysis method which is based on multiple spatial and temporal features. This method utilizes image attention analysis method in the spatial domain. Meanwhile, in the temporal domain, it combines multiple temporal features, which include camera motion, local motion, temporal contrast, and so on. Through the above analysis, we can obtain the spatio-temporal distribution of video attention, which contains saliency map in the spatial domain and attention curve indicating

frame's and shot's attention value in the temporal domain.

Thirdly, based on the video attention analysis method, we propose a spatio-temporal video summarization method. Spatio-temporal summarization means to extract video summary from both spatial and temporal domains, in order to meet the need of video browsing on small screen mobile devices, such as mobile phone. Based on video attention analysis, in the temporal domain, we extract video clips with high attention value as temporal summary. In the spatial domain, to meet the needs of video browsing on small screen devices, we extract regions of interest from the temporal summary as the final spatio-temporal summary. Different with the traditional video summarization methods, this method could extract summary from both temporal and spatial domains. And this method is based on visual attention, which makes the result more accordant with human perception.

Keywords: Video attention analysis, spatio-temporal video summarization.