

学校编码: 10384

分类号 _____ 密级 _____

学号: 23020060153367

UDC _____

厦门大学

博士学位论文

视觉注意计算模型的研究及其应用

Research and Application on Computational Model of
Visual Attention

陈嘉威

指导教师姓名: 周昌乐 教授

专业名称: 人工智能基础

论文提交日期: 2009年4月

论文答辩日期: 2009年6月

学位授予日期: 2009年 月

答辩委员会主席: 沈 敏

评 阅 人: _____

2009年4月

学校编码: 10384
学号: 23020060153367

分类号 _____ 密级 _____
UDC _____

厦 门 大 学

博 士 学 位 论 文

视觉注意计算模型的研究及其应用

Research and Application on Computational Model of
Visual Attention

陈嘉威

指导教师姓名: 周昌乐 教授
专 业 名 称: 人工智能基础
论文提交日期: 2009 年 4 月
论文答辩时间: 2009 年 月
学位授予日期: 2009 年 月

答辩委员会主席: _____
评 阅 人: _____

2009 年 4 月

申请厦门大学博士学位论文

视觉注意计算模型的研究及其应用



作者：陈嘉威

指导老师：周昌乐 教授

专业：人工智能基础

信息科学与技术学院

2009年4月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下取得的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为()课题(组)研究成果，获得()课题(组)经费或实验室的资助，在()实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）：

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交论文（包括纸质版和电子版），允许论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其他方式合理复制。

本学位论文属于：

1、经厦门大学保密委审查核定的保密论文，于 年 月 日解密，解密后适用上述授权。

2、不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经校保密委审定过的，方可打“√”，未经审批均为公开论文。此声明栏不填写的，默认为公开论文，均适用上述授权。）

声明人（签名）：

年 月 日

摘要

视觉注意是人类信息加工中一项重要的心理调节机制,是人类从外界输入的大量信息中选择和保持有用信息,拒绝无用信息的意识活动,是人类视感知过程中高效性和可靠性的保障。视觉注意计算模型的研究,不但有助于探索人类视觉信息处理的工作机理,而且对于解决数据筛选问题和提高计算机的信息处理效率有着重要的意义,在图像分析与图像理解领域、目标检测、信息检索、机器人视觉、视频通讯等领域也有重要的应用价值。

本文对视觉注意机制及其计算方法进行了深入而细致的研究:分析总结了视觉注意机制的认知神经学理论和神经加工机制;以人类视觉加工的生理学理论为依据,紧密结合计算机视觉计算的要求,构建了一个由特征加工、注意集中和注意控制三部分组成的动态视觉注意计算模型的体系结构。提出了一种双通路和层次化的特征加工结构;提出了一种深度特征和运动特征度量方法,以反映场景时空特性对视觉注意的影响;通过 IFNN 神经网络模拟双通路的特征整合过程,实现注意的集中;提出了一种具有注意保持和唤醒功能的注意控制方式。在此基础上,实现了一个基于客体选择的动态视觉注意计算模型。实验表明:本文所提出的计算方法是富有成效的。

本文首先总结了视觉注意机制的认知神经科学及心理学理论,从生物视觉领域找出计算机视觉可借鉴的神经生理学依据,并以之为出发点,寻找认知心理学中注意机制与计算机科学的结合点,构建一个动态视觉注意计算模型的体系结构,并将视觉注意计算划分为特征加工、注意集中和注意控制三个模块,实现认知神经科学与视觉计算的结合。

在特征加工的过程中,解决了特征的选择,特征的显著性度量和特征加工的层次三个方面的问题。在特征选择方面,将特征分为空间特征与非空间特征两类,空间特征的提取是通过提出一种深度特征与运动特征的计算方法来实现的,用以反映场景的时空特性对视觉注意的影响;非空间特征通过提取亮度、颜色和方向特征得到。各类特征的显著性度量依据视觉反差计算实现。根据生物视觉中特征加工层次和功能的差异,空间特征与非空间特征通过模拟 what 和 where 双通路理论进行加工,各类特征的显著图由子特征之间相互竞争和整合得到。

在注意集中方面,以视觉通路理论为指导,通过使用亮度特征、颜色特征和

方向特征等非空间特征来描述物体的感受, 模拟 what 通路的主要功能; 运动特征和深度特征等空间特征用来描述场景的运动和空间信息, 模拟 where 通路的功能。两个通路的整合通过带有自学习和可调节机制的 IFNN 神经网络实现。根据神经网络的脉冲发放时间进行注意焦点的选择, 当两个通路的输入相关联的时候, 神经网络产生最大的增益, 输出单元的脉冲发放时间会比非相关单元的发放时间更短。

在注意控制方面, 根据视觉注意的神经控制特点, 提出了一种动态视觉注意模型的注意控制方式。通过一个唤醒信号来描述视野中新异刺激的强度, 根据唤醒信号的大小开启或屏蔽阈值来控制注意保持与注意唤醒状态的转换。采用注意焦点跟踪算法来实现动态场景的注意保持; 并提出了一种位置增强方法, 以提高新异刺激所在位置的视觉显著性。

本文的研究比较完整地给出了视觉注意计算的思想与方法, 实现了一个适用性较强的动态视觉注意计算模型, 并提高了该计算模型的理论价值和应用价值。实验表明, 本文提出的模型较好地运用了视觉认知规律, 使视觉注意处理结果更加符合人类视觉感知的基本特征。

关键词: 视觉注意; 视觉显著性; 显著区域; 注意焦点; 特征提取

Abstract

Visual attention is regarded as an essential cognitive process of human visual system. Human vision relies on visual attention mechanism to select the relevant parts of scene, on which higher level tasks can be processed. Since information from only a small region of the visual field can progress through the cortical visual hierarchy, visual tasks can be effectively dealt with by limited processing resources. Visual attention models are based on the biological model of visual attention, which mimic the ability of a visual system. Researching on computational model of visual attention is not only helpful in understanding the working mechanism of human visual system, but also has important application in image analysis and understanding, object detection, information retrieval, robot vision, video communication, and etc.

The dissertation addresses the research on visual attention mechanism and its computational methods. The cognitive neuroscience theories and the neural mechanism of visual attention are analyzed. According to the requirement on computer vision, architecture of dynamic visual attention model based on the biophysics and neurophysiology theories of human visual processing is established. This architecture is mainly composed of three parts: feature processing, attentional capture and attentional control. A system of two-pathway based hierarchical feature process is proposed. The extraction of depth features and motion features is realized to measure the third spatial dimension and the time-scale of the complex environments. The integration between the two pathways in brain is simulated by an integrate-and-fire neural network (IFNN), which is employed to compute the focus of attention. The approach of attentional control for dynamic visual attention model is developed to mimic the sustained attention mechanism. According to the theories and techniques, a visual attention model for dynamic scenes based on object selection is implemented.

The cognitive neuroscience theories and neural mechanism of visual attention are analyzed. To meet the requirement of computer image processing by summarizing the latest study of the biological vision, the biological enlightenment for computer vision

is provided. Under the idea of bionics, the architecture of dynamic visual attention model is established, which relates computer image information processing with biophysics and neurophysiology. The model is mainly composed of three parts: feature processing, attentional capture and attentional control.

Three major problems are solved in feature processing: feature selection, saliency computation and the hierarchical processing of feature. On feature selection, the features are classified into two types: spatial features and non-spatial features. The extraction of spatial features which include depth features and motion features are realized to measure the third spatial dimension and the time-scale of complex environments. The non-spatial stimulus features include intensity, color, orientation, and etc. Saliency computation depends on the computation of feature contrast. According to the differences in hierarchy and function of different features, the processing of spatial features and non-spatial features simulates the two pathways processing in brain. The saliency map of each feature is created by the competition or integration of the sub-features.

On attentional capture, the approaches of feature integration and attention focus are developed based on the two-pathway theory. We use the non-spatial features (including intensity, color and orientation) as the perceptual information about object, which are transmitted in "what" pathway. The perception of spatial and motion information related to "where" pathway are presented by depth features and motion features. An integrate-and-fire neural network (IFNN) is employed to simulate the three-way relationship between the two inputs and response to achieve a dynamic and modulatory property. The correlation between two input sets is implemented by the IFNN, which produces a certain amount of gain when two inputs are consistent. In the case that the stimuli from two pathways are correlated, the interspike interval will be shortened. The focus of attention is allocated at the possible target position in the original image after the interspike interval is calculated.

According to neural mechanism of visual attention, the approach of attentional control for dynamic visual attention model is developed. An arousal signal is used to measure the strength of the new stimuli in scenes. The arousal signal is defined as

parameters for sustained attention judgment. A tracking algorithm is proposed to mimic the sustained attention mechanism in dynamic scenes. If the arousal signal cannot over the threshold, the movement of the focus of attention has to be related to the tracking process. If the arousal signal over the threshold, a method of location enhancement is developed to enhance the saliency of the new stimuli.

The computational model of visual attention for dynamic scenes is the target of the research. The theories and approaches are combined to implement the model, which has a strong applicability and important applied value. The experiment results also show that our computational theories and techniques applied in the system are valid and effective.

Keywords: Visual attention; Visual saliency; Salient location; Focus of attention; Feature extraction

目录

第一章 绪论	1
1.1 研究目的与意义.....	1
1.2 国内外研究现状.....	2
1.2.1 视觉注意的加工方法.....	3
1.2.2 视觉注意计算模型的显著性度量.....	5
1.2.3 注意焦点的选择与转移.....	6
1.3 存在的问题.....	7
1.4 研究目标与内容.....	9
1.5 论文的组织结构.....	10
第二章 视觉注意的理论及其加工结构	12
2.1 视觉注意机制.....	12
2.1.1 注意的选择性理论.....	13
2.1.2 特征整合理论.....	14
2.1.3 注意的选择单元.....	16
2.2 初级视觉特征.....	17
2.2.1 初级视觉特征的判别.....	18
2.2.2 初级视觉特征的特点.....	18
2.3 特征加工的神经生理学基础.....	19
2.3.1 视觉加工的神经机制.....	19
2.3.2 视觉加工的层次化结构.....	25
2.4 视觉注意计算模型及其体系结构.....	27
2.4.1 视觉注意计算应注意的问题.....	27
2.4.2 本文模型的体系结构.....	29
2.5 本章小结.....	32
第三章 引入深度特征的视觉显著性度量	33
3.1 视觉显著性.....	33
3.2 图像特征.....	35
3.2.1 颜色特征.....	36
3.2.2 纹理特征.....	37
3.2.3 形状特征.....	38
3.2.4 空间关系特征.....	39
3.3 非空间特征提取.....	40
3.4 深度特征提取.....	41
3.4.1 双目视觉原理.....	42
3.4.2 区域分割和匹配.....	43
3.4.3 视差图的提取.....	44
3.4.4 视差平面的优化.....	45
3.5 特征显著图的计算.....	46
3.6 本章小结.....	49
第四章 引入运动特征的注意计算	50
4.1 基于运动特征的注意计算.....	50
4.2 运动矢量提取.....	52

4.3 运动特征提取.....	52
4.4 运动特征显著图.....	53
4.5 实验结果分析与比较.....	54
4.6 本章小结.....	58
第五章 基于双通路理论的特征整合	59
5.1 注意的集中.....	59
5.2 视觉加工的双通路理论.....	60
5.3 IFNN 神经网络	61
5.3.1 神经网络的结构.....	62
5.3.2 神经网络的活动方程.....	63
5.3.3 神经网络的学习规则.....	64
5.3.4 神经网络的输出结果.....	64
5.4 实验结果分析与比较.....	66
5.4.1 引入深度特征的视觉注意计算.....	66
5.4.2 引入运动特征的视觉注意计算.....	68
5.4.3 模型的抗噪性.....	70
5.4.4 基于客体的选择.....	71
5.5 本章小结.....	73
第六章 视觉注意模型的控制方式	74
6.1 视觉注意的控制方式.....	74
6.1.1 视觉注意控制加工的神经机制.....	74
6.1.2 模型的注意控制方式.....	76
6.2 注意的保持.....	78
6.2.1 唤醒信号.....	78
6.2.2 注意焦点的跟踪.....	78
6.3 注意唤醒与位置增强.....	80
6.4 本章小结.....	81
第七章 总结与展望	82
7.1 本文的主要贡献与创新.....	82
7.2 本文研究的应用价值.....	85
7.3 下一步研究方向.....	87
参考文献	89
致谢	97
攻读博士学位期间发表的论文、参与的工作及获奖情况.....	98

Contents

Chapter 1 Introduction.....	1
1.1 Research Backgrounds.....	1
1.2 The State of the Art.....	2
1.2.1 The Processing Methods of Visual Attention.....	3
1.2.2 The Saliency Computation of Visual Attention Model.....	5
1.2.3 The Selection and Allocation of Visual Attention Model.....	6
1.3 Limitations of Present Research.....	7
1.4 Target and Contents.....	9
1.5 Chapters Outline.....	10
Chapter 2 The Theories and Architecture of Visual Attention Model.....	12
2.1 The Mechanism of Visual Attention.....	12
2.1.1 The Selective Theory of Visual Attention.....	13
2.1.2 The Feature Integration Theory.....	14
2.1.3 The Underlying Units of Attentional Selection.....	16
2.2 Early Visual Features.....	17
2.2.1 The Distinguish of Early Visual Features.....	18
2.2.2 The Characteristic of Early Visual Features.....	18
2.3 The Biophysics and Neurophysiology Theories of Feature Processing.....	19
2.3.1 The Neurophysiology Theories of Feature Processing.....	19
2.3.2 The Hierarchical Processing of Feature.....	25
2.4 The Architecture of Visual Attention Model.....	27
2.4.1 The Attentive Problems of Visual Attention Model.....	27
2.4.2 The Architecture of Our Model.....	29
2.5 Summary.....	32
Chapter 3 The Saliency Computation Including Depth Feature.....	33
3.1 The Vision Saliency.....	33
3.2 The Image Features.....	35
3.2.1 The Color Feature.....	36
3.2.2 The Texture Feature.....	37
3.2.3 The Shap Feature.....	38
3.2.4 The Spatial Relationship Feature.....	39
3.3 Extraction of Non-spatial Features.....	40
3.4 Extraction of Depth Features.....	41
3.4.1 The Theory of Binocular Stereo Vision.....	42
3.4.2 The Color Segmentation and Local Matching.....	43
3.4.3 Extraction of Disparity Plane.....	44
3.4.4 The Optimization Assignment of Disparity Plane.....	45
3.5 The Computation of Saliency Map.....	46
3.6 Summary.....	49
Chapter 4 The Computation of Visual Attention Based on Motion Features.....	50
4.1 Motion Features.....	50
4.2 Extraction of Motion Vectors.....	52

4.3 Extraction of Motion Features	52
4.4 The Motion Saliency Map.....	53
4.5 Experiment and Discussion.....	54
4.6 Summary	57
Chapter 5 Feature Integration Based on The Two-Pathway Theory	59
5.1 Attentional Capture	59
5.2 The Two-Pathway Theory of Feature Processing	60
5.3 IFNN	61
5.3.1 Structure of IFNN	62
5.3.2 Equation of IFNN.....	63
5.3.3 Learning Rules of IFNN.....	64
5.3.4 Output of IFNN	64
5.4 Experiment and Discussion.....	65
5.4.1 Experiment Results of Depth-guided Visual Attention	65
5.4.2 Experiment Results of Motion-guided Visual Attention	68
5.4.3 Experiment Results of Noise Influence.....	70
5.4.4 Object-Based Selection	71
5.5 Summary	73
Chapter 6 The Attentional Control of Visual Attention Model.....	74
6.1 Attentional Control	74
6.1.1 The Neurophysiology Theories of Attentional Control	74
6.1.2 The Attentional Control of The Model.....	76
6.2 Sustained Attention	78
6.2.1 The Arousal Signal.....	78
6.2.2 The Tracking Process of FOA.....	78
6.3 The Attentional Arousal and Location Enhancement	80
6.4 Summary	81
Chapter 7 Conclusion and Future Work.....	82
7.1 Contributions and Innovations of This Dissertation.....	82
7.2 The Applied Value of This Dissertation.....	85
7.3 Future Work	87
References.....	89
Acknowledgements	97
Author's Publications, Participated Projects and Honors.....	98

厦门大学博硕士学位论文摘要库

第一章 绪论

视觉注意(Visual Attention)是人类视觉一项重要的心理调节机制。人类的视觉系统就是利用注意机制,处理大量的视觉信息并及时做出反应。视觉注意计算模型,是通过借用人类视觉的注意机制提出和发展起来的。

我们开展的是有关视觉注意计算方面的研究工作。本章将首先对视觉注意计算研究做一个全面的综述,包括研究的目的、背景和意义、国内外视觉注意计算模型研究进展以及本文的主要研究内容。

1.1 研究目的与意义

本工作是以视觉注意加工机制为研究对象,建立一个有效的视觉注意计算模型,在计算机图像信息处理中模拟和实现人的视觉注意机制,并将其具体应用于动态场景的感兴趣区域检测之中。

视觉注意计算是一个多学科交叉的研究领域,本研究的目的是通过将传统的图像信息处理过程和人类视觉加工机制相结合,形成一个合理的计算资源分配方案,引导整个图像信息处理过程,快速有效的获取那些容易引起观察者注意的显著区域,使计算机系统的图像信息处理具备类似人的视觉主动性和选择性。

视觉注意计算研究的实际意义是由视觉注意机制在人类视觉系统中的重要性及其科学价值决定的:

(1) 视觉注意机制是人类信息加工过程中的一项重要心理调节机制

视觉是人类最主要的信息来源。在大脑接收的来自外部世界的大量信息中,绝大部分是通过视觉系统进行加工和处理的。人类视觉系统在处理外界信息的时候,具有非常突出的数据筛选能力。人所处的周围环境及人本身的时间空间位置都处在不断运动的过程中,这客观地构成了人脑在不断地接受来自内部和外部的巨大信息刺激。同时,觉醒状态的人在不断的进行着心理活动和信息加工,这就意味着人必须不断地从来自内部和外部的巨大信息中选择并保持有用信息,拒绝无用信息,以便迅速地从中筛选出那些与自身相关的重要信息,并及时做出反应,使其拟进行的心理活动不受干扰地进行下去。这种具有选择性和主动性的保

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士学位论文摘要库