

基于信息融合的阅读过程识别

刘英梅

(肇庆学院图书馆, 广东 肇庆 526061)

〔摘要〕本文在分析阅读过程与信息融合的基础上, 提出了基于信息融合的读者阅读过程识别的新思路。重点构建了阅读过程信息融合功能模型, 介绍了融合算法、阅读过程融合分类、阅读信息获取的一些方法。

〔关键词〕信息融合; 阅读过程; 识别

DOI: 10.3969/j.issn.1008-0821.2012.01.012

〔中图分类号〕G203 〔文献标识码〕A 〔文章编号〕1008-0821(2012)01-0052-04

Reading Process Recognition Based on Information Fusion

Liu Yingmei

(Library, Zhaoqing University, Zhaoqing 526061, China)

〔Abstract〕On the basis of analyzing reading process and information fusion, the new viewpoint of reading process recognition based on information fusion was proposed in the paper. Reading process information fusion function model was constructed, and some methods of fusion algorithm, reading process fusion and classification, reading information acquisition were introduced.

〔Key words〕information fusion; reading process; recognition

信息融合一词首先来源于美国, 它的研究起源于军事系统建设的需求。20 世纪 90 年代, 随着智能机器人、图像处理和计算机多媒体技术的发展, 迎来了信息融合研究的世界性热潮^[1]。信息融合的概念在不同的学科有不同的表述, 在信息智能系统研究领域, 信息融合是由多种信息源(如传感器、数据库、知识库和人类本身等)获取信息, 并进行滤波、相关和集成, 从而形成一个适合信息选择达到统一目的(如目标识别跟踪、传感器管理和系统控制等)的表示构架^[2]。

阅读活动是读者从记录知识的载体中获得信息与知识的社会实践过程、生理过程与心理过程^[3]。读者阅读过程中的特征参数, 表现为非线性、时变性、复杂性、多样性和随意性。对阅读过程中的知识载体信息、人体表面肌电信息以及运动图像解析等有关参数需要借助多种数学工具, 进行了全方位的深入研究, 以期提高对阅读过程的精准解析, 读者阅读过程识别为研究读者心理、指导阅读和合理构建知识载体采购系统打下基础。

1 信息融合的处理过程

按照融合过程中信息抽象的层次, 可以将信息融合过程分为: 数据融合、特征融合、融合决策 3 个层次^[3]。数据层融合直接对原始数据进行配准和关联, 特征层融合对特征向量进行配准和关联, 然后再进行识别, 而决策层融合则是先进行识别, 再对各个决策结果进行关联, 得到融合的判决结果。

输入 → 数据融合 → 特征融合 → 融合决策 → 输出

图 1 信息融合的处理过程

信息融合处理总是面向具体应用的, 针对一个确定的具体决策任务设计融合处理流程。图 2 是针对阅读过程的信息融合功能模型。

从图 2 可以看出, 整个阅读融合处理过程是以阅读信息的获取与处理为基础的, 无论是阅读模型的建立还是阅读状态的评判, 最终都是基于对阅读过程中所获得的多源信息综合分析的基础之上的, 因此, 阅读信息获取是基础。

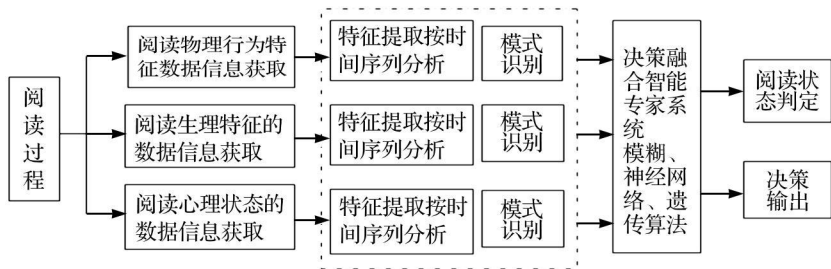


图 2 针对阅读过程的信息融合功能模型

2 阅读过程分析

阅读是指从文字、声音、图像及其三位一体的表达方式中提取信息的过程。也就是说，通过视觉、听觉器官接受文字、声音、图像的信息，再经过大脑的编码加工，从而理解阅读信息的意义。对阅读的研究是一个广泛而复杂的领域^[4]。

阅读过程可以分为 3 个层次：

(1) 物理学层次：阅读过程中的行为姿态和面部表情，包括身体的移动、翻书的动作、鼠标的移动以及眼部、面部和口部肌肉的变化等。

(2) 生理学层次：读者阅读过程中的神经冲动，如人的皮质醇水平、心率、血压、呼吸、皮肤电活动、掌汗、瞳孔直径等。

(3) 心理学层次：阅读过程中读者对信息的接受、编码、贮存、提取和使用，这一过程可以归纳为 4 种模式：即感知、记忆、控制和反应。读者已有的知识和知识结构对其行为和当前的阅读活动起决定作用。

2.1 阅读过程模式识别方法

阅读过程模式识别就是根据阅读者和阅读对象的某些特征把它们归到一定的类别^[5]，是对感知数据的解释过程。

阅读过程行为识别大体可以划分为 3 层：从阅读背景中发现目标，通常称为目标检测；区分阅读目标的类型，通常称为目标分类；个体目标的确认，通常称为目标确认。这 3 个层次因其分类任务和内涵的不同，对目标表述的精细程度的需求不一样。3 个层次是互相联系的，目标检测是基础的层次，目标分类一般建立在目标检测的基础上，对目标表述的精细程度需求要求高些，目标确认一般建立在目标分类层次上，对目标表述的精细程度需求更高。

2.2 阅读过程行为特性分析

阅读过程行为特性分析是目标识别的基础，包括不同类别目标及其背景的特征属性和影响因素。测量和分析目标特性需要考虑目标可能存在的变化模式，相关背景的变化模式和传感器感知过程会出现的多种模式等。

2.3 阅读过程行为识别模型的建立

阅读过程行为模型的建立是目标识别的关键部分。目标识别模型有 3 种方法：第一种是目标模板表述模型，直接使用目标识别相关的典型目标数据建立模型，采用相似

性度量的模板匹配进行识别；第二种是基于数据层的特征矢量表述模型，使用单层结构的多特征属性组合建立模型，在特征矢量表述模型基础上可以建立形形色色的识别分类方法；第三种是基于元特征的层次结构表述。通过对目标数据的分析，可以提取以基础单元的特征，还可以将相连的具有同样属性的特征集成特征集，即形成所谓的元特征。建立目标识别模型从本质上说，是用智能计算过程来实现人们识别处理过程。

3 融合算法

早在 20 世纪 80 年代，Hall 和 Llinas 考察了 30 个信息融合系统，并归纳了使用的 75 种融合算法。目前信息融合算法有上百种，按概念分类，目标识别的融合算法主要有：物理模型算法、参数分类算法、基于认识模型的算法^[6]。

3.1 物理模型算法

这种算法根据物理模型直接计算实体特征。预测一个实体特征的物理模型必须以被识别物体的物理特征为基础，而实际物理模型往往相当复杂，建立起来非常困难。尽管实际中很少使用这种方法，但在基础研究工作中却使用它。常用的技术包括仿真（Simulation）、估计以及依照句法（Syntactic）的方法。其中估计方法有卡尔曼滤波、最大似然和最小均方估计等。

3.2 参数分类算法

这种算法寻求一个标识说明使之依赖于参数数据而不是物理模型，在参数数据和一个标识说明间建立一个直接的映像。这种方法可进一步分为基于统计的算法和基于信息的算法^[7]。

基于统计的算法：统计算法包括：经典推理、Bayes 推理和 D-S 证据理论。

基于信息的算法：依赖于观测参数与目标身份之间的映射关系来对目标进行标识。信息论方法包括参数化模板、人工神经网络、聚类算法、优化图解（Figures of Merit）和相关性度量。

3.3 基于认识模型的算法

基于认识模型的信息融合算法是通过模仿人类辨别实体的识别过程模型，解决证据不确定性或决策中的不准确性问题。也就是说，信息融合过程中存在的不确定性可以直接用模糊逻辑表示，然后使用多值逻辑推理，根据各

种模糊演算对各种命题（即各传感器提供的数据）进行合并，从而实现信息融合^[7]。基于认识模型的算法主要包括逻辑模板、知识（专家）系统和模糊集合论。

4 读者阅读过程的融合分类

数据分类是将原始感知的数据划分为具有相同特征属性的不同子集，也就是赋予其具有一定语义特性的标记。读者阅读过程的融合分类分为数据层融合分类、特征层融合分类和决策层融合分类。由于成像传感器的应用，以图像为中心的融合分类研究应用比较广泛。

4.1 数据层融合分类

基本处理方法就是多源图像数据的直接加权融合和多源图像的变换融合。前者通过对多源图像数据相对位置的图像灰度值加权形成一幅新的融合图像，以后在这基础上进行分类。后者是将多源图像数据变换成一个新的图像集，常见方法有：变换为彩色图像进行分类，经过 PCA 变换进行分类，经过 ICA 变换进行分类。

4.2 特征层融合分类

- (1) 多源信息的特征提取，如线段、区域、轮廓以及指纹、人脸、声音等。特征的提取和具体应用与使用的传感器有关。
- (2) 分类器设计与选择，如距离、神经网络方法、SVM 等分类器形式和决策树的层次处理结构等。
- (3) 应用特征分类的其他形式，如采用心理学中视觉竞争机理的图像显著性局部区域的快速提取。

4.3 决策层融合分类

决策层融合分类在很大程度上是指多分类器融合分类。根据实际应用常常使用不同类型的传感器，如语音和图像融合识别说话人的系统，其数据具有固有的不同步、不完全对应等特点，采用特征层融合会遇到一些困难，采用决策层融合可以缓解。

5 阅读信息获取

在阅读信息获取过程中，常常面对不同感知对象、不同类型的平台和不同类型传感器等情况相应的融合处理需求也不尽相同。表 1 是阅读过程中的信息融合的不同需求。

表 1 阅读过程中的信息融合的不同需求		
对 象	融合处理	目 的
单(多)目标	平台	阅读什么，在哪儿阅读
多目标	处理中心	什么时候需要信息，需要什么信息

在阅读信息的获取中，一方面要充分考虑可能获得的所有信息资源，选择在阅读范围和阅读特征上彼此互补的信息资源，以满足预定决策任务的需求；另一方面尽可能减少甚至消除获取的信息资源的不确定性和不可靠性。

阅读过程中目标的属性常常是多种多样的。有些属性可以直接表现出来并被传感器直接感知，有些则是非直接表现出来的，可以通过感知的数据经过变换和处理获得。当采用多传感器和传感器网络感知时，采集或搜集有关数据时，必须考虑到传感器实际受到的 3 个基本方面的约束，即感知时空域的限制，感知特性的限制和工作环境与条件的限制。下面是已知的阅读过程中读者数据获取数据方式：

- (1) 将传感器（如摄像头）安装在读者阅读活动场景中，由传感器采集读者的阅读活动图像序列，通过对图像或视频处理实现阅读活动分析。这种方式是获取阅读活动信息的主流方法。当然，从复杂背景中提取人体阅读活动信息难度大，而且存在图像分析的算法复杂，难以实时处理的问题^[8]。
- (2) 通过三维跟踪技术获得读者的运动位置，然后进行运动分析。常用的三维跟踪技术有光学跟踪、电磁跟踪和声学跟踪等。中国科学院计算技术研究所^[9]在基于动态背景构造的视频运动人体提取、基于视频的人体运动跟踪、三维人体运动模拟与仿真以及基于运动库检索和视频分析的三维人体运动参数获取等方面攻克了多项关键技术。主要有：运动人体提取技术、运动人体跟踪技术、三维人体运动模拟与仿真技术、三维人体运动参数获取技术。
- (3) 无损伤脑成像技术，记录大脑对屏幕、文字和图像的反应，能够进行读者阅读行为的科学研究。脑内存在许多功能分离而又相互协同工作的脑区，大脑对阅读信息的认知和情绪反应很多时候是在无意识情况下发生的，人在进行不同的心理活动时，脑内有不同的脑区被激活^[10]。Gallagher 等^[11]利用功能磁共振成像（functional Magnetic Resonance Imaging, fMRI）技术揭开心理认知的神经生物学基础，因此这种迅速发展的探测人脑如何加工获取信息并作出决策的能力，为科学地研究阅读过程提供了依据。
- (4) 事件相关电位（Event-related Potential, ERP）是脑波的一种，将记录到脑电图的脑部原始生理信号进行再分析处理而得，为了解人脑的高级认知活动打开了一扇窗口。湖南师范大学的陈竹使用事件相关电位方法对不同个性大学生阅读攻击性图片的研究显示出大脑皮层的不同反应，可以提供研究人的认知活动和判断人的认知能力的有用指标^[12]。
- (5) Raichle^[13]介绍了认知神经科学家用来研究认知过程的正电子发射层析摄影扫描技术（PositronEmission Tomography, PET），这是一种基于减法逻辑的方法，可观察由纯实验条件所引起的脑兴奋区域和兴奋水平。PET 曾多次被用来研究情景记忆（Epi-sodic Memory），当被试者提取情景记忆时，右前额叶显示较其它记忆提取时更高的兴奋水平^[14]。
- (6) 株式会社岛津制作所的河野理等发明的一种应用核磁共振现象的磁共振成像装置，其特征在于，它包含，

阅读倾斜场脉冲的产生手段装置，通过所述第三倾斜磁场线圈产生所述阅读倾斜磁场脉冲，它与所述回波信号成一时间关系，这种装置当采用混合扫描方法摄取图像时，能够修正回波信号之间的相位差，从而在不加重控制系统负荷的情况下抑制图像模糊^[15]。

华南师范大学的金花等应用功能性磁共振成像技术实验发现：预期段落的阅读任务激活——特异性皮层网络^[16]。

(7) 美国 MIT 媒体实验室已经开发出数种用于测量心理信号的传感器^[17-19]，如用于测量皮肤电信号的皮肤电反应传感器、血流脉冲传感器、肌动电流图传感器、呼吸传感器等。压力感应鼠标^[20]和自动面部表情分析^[21]也作为非语音手段，对人类的情感状态进行识别。

(8) 哈尔滨工业大学的金辉和中国科学院的高文^[22]提出一种人脸面部混合表情识别系统。首先把人脸检测定位，通过图像预处理分别提取其运动特征，并按时序组成特征序列，然后分析不同特征区域所包含的不同表情信息的含义和表情的含量；最后通过信息融合来理解、识别任意时序长度的、复杂的混合表情图像序列。其各种表情的总体识别率达到 96.9%。

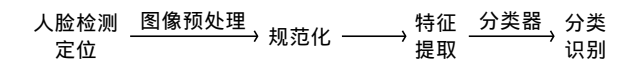


图 3 人脸识别算法流程

(9) 北京理工大学的续爽等^[23]在图嵌入的框架下提出一种根据表情相似度构建邻接权重图的方法来学习人脸表情子空间。将人脸分为 6 种表情，建立一种人脸表情流形模型，对各种可能的人脸表情图像进行全局表示，从数据集中各种可能的人脸表情数据中提取本质的表情特征，所提出的子空间方法很好地解决了个体、光照、姿态等差异对人脸表情识别带来的困难。

基于信息融合的阅读过程识别就是从不同传感器的角度观察待识别目标，对待识别目标做出预处理、特征提取、识别或判决，通过在不同层次将多传感器所观察的信息进行融合最终获得联合推断结果，以达到提高识别效果的目的。这对于研究读者阅读的认知，具有重要意义。

参 考 文 献

[1] 边肇祺, 张学工, 等. 模式识别 (第二版) [M]. 北京: 清华大学出版社, 2000.
[2] 权太范. 信息融合神经网络——模糊推理理论与应用 [M]. 北京: 国防工业出版社, 2002.
[3] 王润生. 信息融合 [M]. 北京: 科学出版社, 2007.
[4] 祖林, 石继业. 阅读过程心理机制的研究概述 [J]. 中国电力教育, 2008, (2): 179- 181.

[5] 中国大百科全书编辑委员会. 中国大百科全书·教育卷 [Z]. 北京: 中国大百科全书出版社, 1985.
[6] 胡勇. 信息融合在模式识别中的应用研究 [D]. 合肥: 合肥工业大学, 2004, (3).
[7] 涂小强. 信息融合的原理与方法概述 [J]. 电讯技术, 1999, (3): 1- 6.
[8] 雷建和. 基于多源信息融合的人体运动分析与建模研究 [D]. 合肥: 中国科学技术大学, 2006, (5).
[9] 中国科学院计算技术研究所. 基于三维人体运动仿真与视频分析的计算机辅助运动系统及示范应用 [Z]. 北京: 中国科学院计算技术研究所, 2008.
[10] 王志良, 郑思仪, 王先梅, 等. 心理认知计算的研究现状及发展趋势 [J]. 模式识别与人工智能, 2011, (2): 215- 225.
[11] Gallagher H, Happe F, Fletcher P. Reading the Mind in Cartoons and Stories: An fMRI Study of ‘Theory of Mind’ in Verbal and Non-verbal Tasks [J]. Neuropsychologia, 2000, 38 (1): 11- 21.
[12] 陈竹. 不同个性大学生阅读攻击性图片的 REp 研究 [D]. 湖南: 湖南师范大学, 2006.
[13] Raichle M E. Visualizing the Mind [J]. Scientific American, 1994, 270 (4): 36- 43.
[14] Wheeler M, Stuss D, Tulving E. Toward a Theory of Episodic Memory: The Frontal Lobes and Autonoetic Consciousness [J]. Psychological Bulletin, 1997, 121 (3): 331- 354.
[15] 河野理, 等. 磁共振成像装置 [P]. 日本京都市: CN1104883, 1995- 07- 12.
[16] 金花. 文本阅读中预期推理生成的脑激活模式 [A]. 上海: 第十届全国心理学学术大会论文摘要集 [C]. 2005.
[17] Wilson G M, Sasse M A. From Doing to Being: Getting Closer to the User Experience [J]. Interacting with Computers, 2004, 16 (4): 697- 705.
[18] Picard R W, Scheirer J. The Galvactivator: A Glove that Senses and Communicates Skin Conductivity [Z]. Proc of the International Conference on Human- Computer Interaction. New Orleans, USA, 2001: 91- 101.
[19] Scheirer J, Fernandez R, Klein J, et al. Frustrating the User on Purpose: A Step toward Building an Affective Computer [J]. Interacting with Computers, 2002, 14 (2): 93- 118.
[20] Qi Yuan, Picard R W. Context- Sensitive Bayesian Classifiers and Application to Mouse Pressure Pattern Classification [Z]. Proc of the 16th International Conference on Pattern Recognition. Quebec, Canada, 2002, VIII: 448- 451.
[21] Sebe N, Lew M S, Sun Y, et al. Authentic Facial Expression Analysis [J]. Image and Vision Computing, 2007, 25 (12): 1856- 1863.
[22] 金辉, 高文. 人脸面部混合表情识别系统 [J]. 计算机学报, 2000, 23 (6): 602- 608.
[23] 续爽, 贾云得. 基于表情相似性的人脸表情流形 [J]. 软件学报, 2009, 20 (8): 2191- 2198.