

基于事例的面部动作单元识别算法^{*})

王上飞 薛 佳 王煦法

(中国科学技术大学计算机科学与技术系 合肥 230027)

摘 要 本文提出基于事例的交互式遗传算法进行面部动作单元识别的算法,将用户的比较能力融入到搜索过程,快速检索到与待识别图像匹配的事例图像,从而实现动作单元的半自动识别。该方法不需抽取图像特征,因而可用于识别非控制成像条件下自发面部图像或图像序列中的动作单元,具有较好的鲁棒性和实用性。文中采用 16 幅受控成像条件下收集的简单图像进行实验,单独 AU 的平均识别率达到 77.5%,AU 组合的平均相似度为 82.8%。采用 10 幅有干扰的非受控成像条件下收集的复杂图像进行实验,单独 AU 的平均识别率为 82.8%,AU 组合的平均相似度为 93.1%。相对于特征脸算法,本文算法的平均识别率和相似度都有较大幅度的提高。

关键词 交互式遗传算法,面部动作单元,自动识别

Algorithm of Case-based Facial Action Units Recognition

WANG Shang-Fei XUE Jia WANG Xu-Fa

(Department of Computer Science and Technology, University of Science and Technology of China, Hefei 230027)

Abstract This paper proposes a case-based automatic facial action units recognition approach using interactive genetic algorithm, which embeds human compare ability into search process. To obtain AU codes of a new facial image, IGA is applied to retrieve a match instance to this new one from the case base based on users' evaluation. Then the solution suggested by the matching case is used as the AU codes to the new facial images. The approach is easy and has the latent capacity to detect AU codes from facial images and image sequences insensitive to pose and occlusions, since feature extraction is not needed. The effectiveness of our approach is evaluated by 16 standard images and 10 un-standard images. A recognition rate of 77.5% is achieved on single AUs, and a similarity rate of 82.8% is obtained on AU combinations to standard images. A recognition rate of 82.8% is achieved on single AUs, and a similarity rate of 93.1% is obtained on AU combinations to un-standard images. The compare experiments with eigenface method also demonstrate the effectiveness of our approach.

Keywords Interactive genetic algorithm, Facial action unit, Automatic recognition

1 引言

在人与人交流的过程中,识别和理解对方的情感状态,并做出相应反应,是人类情感智能的核心^[1],也是人类整个智能中不可或缺的部分。近年来兴起的情感计算(Affective Computing, AC)的研究,旨在赋予机器情感智能^[2],其研究内容包括情感的识别、建模、理解和表达等。由于面部表情是人类情感交流的一个重要途径,因此面部图像或图像序列中的情感的自动识别研究已经成为情感计算中的一个重要课题。

由美国心理学家 Ekman 和 Friesen 提出的面部动作编码系统(Facial Action Coding System, FACS),是行为科学和情感研究中针对面部表情使用最广泛的标准^[3]。FACS 将面部表情分成 44 个动作单元(Action Units, AU)以及一些头部和眼部的位置和动作描述,这些动作单元与面部的肌肉活动相关联。研究表明,在自发产生的面部动作中有超过 7000 个面部 AU 组合。

当前,FACS 编码主要是由编码专家编写。编码专家需要经过特殊的学习和训练。这样的方式效率低,速度慢。自

动生成 FACS 编码对于面部表情的自动识别与理解研究有着革命性的影响。但是,利用计算机自动识别 AU 仍然很困难。

目前已经有一些关于 AU 自动识别的研究^[4~9],这些研究的主要框架是:首先抽取描述面部特征参数,然后使用机器学习技术识别 AU。为了使面部特征检测更加可行,大多数研究中所使用的数据库是在被控制的成像环境下收集的,如没有人脸附属物(如首饰、眼睛、帽子等)、恒定的照明条件、被观测者做出指定的表情、正面或以特定的角度面对摄像机等(这类图像本文称为简单图像,而在非控制的成像环境下收集的图像本文称为复杂图像)。这些特定的成像环境在实际应用中难以满足,因而成为将自动识别系统投入实际应用的一个主要障碍。

针对这一问题,本文提出基于事例的面部动作单元识别算法(Algorithm of Case-based Facial Action Units Recognition)。该方法不需提取面部特征参数,直接采用交互式遗传算法(Interactive Genetic Algorithm, IGA)识别 AU^[10],因而减少了对图像成像环境的要求,具有更好的实用性。以往研究中,已知 AU 的训练集数据作为本算法的事例集,事例图像

^{*}) 本文受国家自然科学基金—人工情感智能模型及其应用研究(NO. 60401004)资助。王上飞 博士,主要研究方向:情感信息处理,计算智能,信息检索;薛 佳 硕士研究生,主要研究方向:情感信息处理;王煦法 教授,博导,主要研究领域:计算智能、情感信息处理、人工生命。

的 AU 编码作为解决方案。为了获得一个新的面部图像的 AU 编码,算法采用 IGA 从事例集中检索一个最接近的事例图像,该图像的 AU 编码即成为新图像的 AU,从而实现 AU 的自动识别。该方法简单,且不受姿势、面部附属物等成像条件的影响。为了验证算法的有效性,本文进行了两类实验,第一类实验对 Cohn-Kanade 面部表情数据库中的 16 幅在特定成像环境下拍摄的简单图像进行 AU 识别,第二类实验针对网上收集的 10 幅有较多干扰的复杂图像进行 AU 识别。获得了较好的实验结果。另外,将本文的算法与特征脸方法进行比较,本文的算法具有较高的平均识别率和平均相似度。

本文组织如下:第 1 部分引言,介绍相关的研究背景;第 2 部分详细阐述本文的基本思想和方法;第 3 部分为实验结果及其分析;最后是结束语。

2 基于事例的面部动作单元识别

2.1 交互式遗传算法

IGA 是针对一些应用领域存在的适应度函数难以明确表达的问题,通过用户参与进化过程,对进化个体进行评估以代替计算过程的一种改进方法,是一种基于用户主观评价的全局搜索优化技术^[11]。换言之,就是在进化的过程中,采用人机交互的方式,由用户确定适应度函数的值,通过选择、交叉、变异等遗传操作在参数空间进行搜索。图 1 是一个典型的 IGA 系统。用户观察或者聆听系统的输出,并且根据系统输出和用户目标之间的距离来给出主观评价,即适应度。GA 通过用户的评价值优化目标系统从而获得最优的输出。这样,IGA 就成为一个综合了人类感觉、知觉、认知能力和 GA 搜索能力的技术。在过去 10 年中,IGA 已经被广泛应用到计算机图像学、设计、信号处理、数据挖掘、数据库检索等方面。本文把 IGA 的应用扩展到一个新的领域:从面部图像中识别 AU。

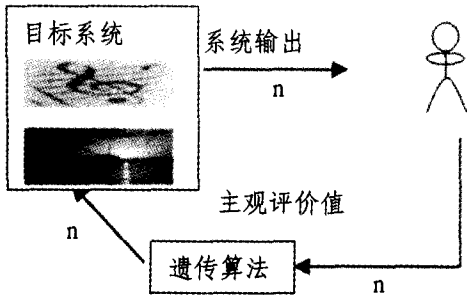


图 1 IGA 基本框架:基于主观评价值的系统优化技术

2.2 基本思想

以前的 AU 识别研究遵循一个两阶段框架,即首先抽取特征,使用整体空间分析或局部空间分析等方法。其次,使用机器学习技术,如神经网络,通过对训练样本的学习获取特征和 AU 之间的映射函数,从而实现 AU 的自动识别。虽然这个框架可以获得一些好的结果,但是特征抽取需要一些受控制的成像环境,它限制了这些研究的实际应用。本文提出的面部 AU 识别方法没有使用特征抽取,而是直接使用 IGA。

虽然没有受过任何训练的实验者不能从面部图像中检测到 AU,但是他们可以很容易识别诸如嘴巴张开、皱眉、瞪眼等面部行为,并能轻易地把握两幅图像在面部行为上的相似

点和不同点。即使图像中的人物不是同一个人,或者有不同的面部附属物,这些面部行为在解剖学上与面部肌肉的收缩相关,而这些面部肌肉的运动又直接与 AU 的定义相关。

本文的基本思想是把人类的比较能力嵌入到 GA 中,从而实现 AU 的自动识别。GA 在图像事例库中搜索并且把检索结果显示给用户,用户根据它们与新的面部图像之间的相似程度给出它们的适应度。这个检索过程一直重复,直至用户找到最相似的事例图像。然后这幅事例图像的 AU 编码就用作新图像的 AU 编码。

2.3 面部动作单元识别算法

图 2 给出基于事例的 AU 识别算法框架。

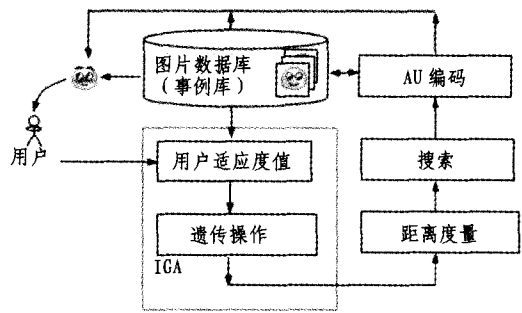


图 2 基于事例的面部动作单元识别算法框架

这里,事例库中的面部图像是已知 AU 编码的图像,这些在其他研究中作为训练数据集。事例图像的 AU 编码作为相应的解决方案。为了识别新图像中的 AU,系统从事例库中随机选择出 N 个事例图像作为初始群体显示给用户。用户观察这些事例图像和新图像的面部动作的相似程度,给出其适应度。这里适应度分为 7 级。越相似的图像,适应度值越大,反之亦然。根据用户的评价值,系统通过遗传操作产生下一代事例图像,并且显示给用户。用户再对显示的图像给予评价。不断重复,直至用户找到最接近新图像的事例图像或达到最大的迭代次数。然后匹配事例的解决方案就作为新的面部图像的 AU 编码。

2.2.1 遗传编码描述

事例图像的 AU 编码转化为 92 位二进制串($J = 92$),如表 1 所示。编码中包含最常用的 AU 及其位置信息,左、右、上和下。在本文的编码系统中暂时没有考虑 AU 的强度。

表 1 遗传编码描述

AU1	AU2	L	R	AU4	L	R	...
-----	-----	---	---	-----	---	---	-----

2.2.2 遗传操作

这里采用赌轮选择、多点交叉和随机变异,每一代产生 N 个候选项。

2.2.3 距离度量

获得 N 个候选项之后,按式(1)计算出事例库中所有图像与候选项之间的距离,取出最小具有距离的 N 个事例图像作为下一代的个体,显示给用户,供用户做进一步分析判断。

$$dist(j,i) = \sum_{k=1}^n (FI_j(k) \oplus FC_i(k)) \tag{1}$$

式(1)中, $FI_j(k)$ 表示第 j 幅事例图像的第 k 个位置的 AU 编码值, $FC_i(k)$ 是第 i 个候选项的第 k 个位置的遗传编码。

3 实验结果与分析

本文的实验是在 Cohn-Kanade 面部表情数据库的基础上进行的。该数据库含有超过 200 个被观察者的近 2000 个面部图像序列以及 481 幅图像所对应的 AU 编码。这 481 幅已知 AU 编码的图像构成本文实验的图像库。

实验者是 5 位年龄在 22~29 岁之间的本校老师和学生,其中 3 名男性、2 名女性。实验者中,一位参与本系统的设计,因而了解 AU 的基本知识。其余各位没有任何关于面部表情度量或者是本系统的知识背景。所有的实验者都不是 AU 编码专家。表 2 给出实验的 IGA 参数。

表 2 IGA 参数

参数名称	参数值
种群大小(N)	12
编码长度(J)	92
交叉概率	0.8
变异概率	0.01

3.1 实验一:简单图像的 AU 识别实验

该次实验中,我们从 Cohn-Kanade 面部表情数据库中已知 AU 编码的 481 幅图集中抽出 16 幅作为待识别 AU 的测试图像,其余 465 幅作为事例库。Cohn-Kanade 面部表情数据库是在特定环境下收集的、干扰较少的简单图像库。图 3 为系统的界面。

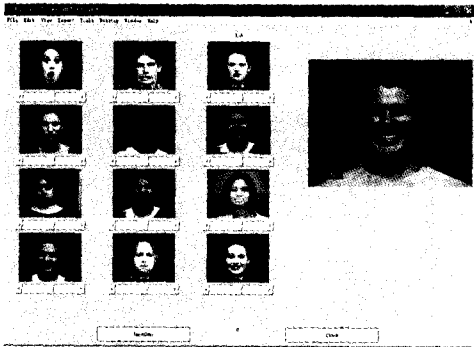


图 3 基于事例的面部动作识别系统界面

表 3 单独 AU 的实验结果

AU 号码	平均识别率
1	84.0%
2	95.0%
4	70.0%
5	90.0%
6	55.0%
7	46.7%
9	80.0%
12	100.0%
15	90.0%
17	100.0%
25	80.0%
26	66.7%
27	50.0%

实验者比较测试图像和其它 12 幅系统显示的事例图像,并根据它们之间的相似度给出 12 幅事例图像的适应度。系

统依据用户的评价值,通过遗传操作产生下一代事例图像,显示给用户。不断重复,直至用户找到最相似的事例图像。在该次实验中,每个实验者平均需要大约 2min 完成一幅测试图像的实验。表 3 和表 4 列出 80 人次(16×5)实验的统计结果。

由表 3 可知,单个 AU 的总平均识别率是 77.5%。有些 AU,如 AU2(外部眉毛抬升)、AU12(唇角抬升)、AU17(下巴抬升)的平均识别率较高,而 AU6(脸颊抬升)、AU7(眼睑收紧)、AU27(嘴部伸展)的平均识别率较低。我们认为这些平均识别率较低的 AU 可能在动态图像中较容易被识别出来,而实验系统使用的是静态图像。其次,也可能与人类视觉的重点有关系。

表 4 AU 组合的实验结果

图像号码	AU 组合	平均交互次数	平均相似度
1	1+2+5+27	3.40	75.0%
2	4+17	3.40	70.0%
3	1+2+27	2.80	86.7%
4	1+4+15+17	3.00	75.0%
5	6+12+25	2.40	100.0%
6	4+7+9+17	2.60	80.0%
7	1+2+26	2.60	80.0%
8	15+17	2.00	90.0%
9	1+2+5+26	2.40	75.0%
10	12+25	2.60	70.0%
11	4+6+7+9+17	3.00	68.0%
12	6+12	3.20	80.0%
13	25	2.80	100.0%
14	26	2.80	100.0%
15	6+7+12+25	2.40	75.0%
16	12	2.60	100.0%

表 4 列出每个测试图像与检索到的匹配事例图像的平均相似度和平均交互次数。这里我们定义相似度为测试图像和匹配事例共用的 AU 数目与测试图像的 AU 数目之比值。由表 4 可知平均交互次数是 2.8 次,平均相似度是 82.8%。这意味着通常经过 3 次交互之后就可以找到一个非常相似的事例图像。测试图像和匹配事例之间的平均相似度超过 80%。

3.2 实验二:复杂图像的 AU 识别实验

3.1 节中使用 Cohn-Kanade 面部表情数据库测试本文提出的方法,该数据库是在受控制的成像环境下收集的,因此在一定程度上说明了方法的可行性。为了验证本文方法的普遍适用性,本节实验使用非控制成像环境下收集的复杂面部图像。实验图像来源于网上的新闻人物和人物写真。允许图像中包含对前人方法中的特征抽取产生严重影响的因素,如眼镜、留海、帽子、手遮住局部脸部、复杂背景、观察角度不正、光线不强等。测试图像如图 4 所示。实验界面与过程和 3.1 节中的类似。实验中的事例集仍是 Cohn-Kanade 面部表情数据库。

在本实验中,每个实验者平均需要使用 2.5min 的时间完成一幅测试图像的实验,这里的实验时间比 3.1 节的时间稍微多的原因是测试图像不像前次实验里面的测试图像一样是跟事例图像一种样式,因此需要稍微多一点时间进行熟悉,而且这里的测试图像可能有各种各样的多余信息,这需要实验

者排除这些多余信息。但是总体来说,增加的时间并不是很多。



图 4 10 张测试图像

表 5 和表 6 给出本次实验的 50 (10×5) 人次实验结果的统计分析。

表 5 单独 AU 的实验结果

AU 号码	平均识别率
4	100.0%
6	85.7%
7	60.0%
9	40.0%
12	97.1%
17	100.0%
25	97.1%

表 6 AU 组合的实验结果

图像号码	AU 组合	平均交互次数	平均相似度
1	6+12+25	2.40	100.0%
2	6+12+25	2.60	100.0%
3	6+12+25	2.60	100.0%
4	6+12+25	2.40	100.0%
5	25	2.60	100.0%
6	4+6+7+9+17	2.60	68.0%
7	12	2.40	100.0%
8	6+12	2.20	70.0%
9	6+12+25	2.40	93.3%
10	25	2.80	100.0%

由表 5 中可知,对于单个 AU 的总平均识别率是 82.8%。表 5 中,AU9(鼻部起皱)的平均识别率偏低,而在表 3 中却没有出现这样的情况。经分析,实验 2 中的 AU9 是在第 6 幅图像中出现的,而第 6 幅图中有两个人,我们的测试对象是孩子的脸。在显示的时候,图像较小,实验者不易分辨,需要更好的显示界面。

表 6 中列出了 AU 组合平均相似度和平均交互次数。这里的相似度定义和实验 1 的定义相同。计算得出总平均相似度是 93.1%,总平均交互次数是 2.53 次。也就是说,经过大约 3 次交互就可以得到与测试图像很相似的事例图像,且测试图像是有很多干扰因素。因此可以说本算法是可行的。

3.3 比较实验

特征脸方法是 AU 识别的常用方法之一,要求面部图像中各个器官的位置相对固定,因此这里只对简单图像采用特征脸方法进行 AU 识别。

实验中首先对 Cohn-Kanade 图像库中的图像手工进行标准化,标准化后图像的尺寸大小为 260×360 像素,为眼睛位置固定的灰度图像。训练集和测试集分别与实验 1 中的事例集和测试集相同。通过主元分析找到训练集图像库的特征脸空间,计算出训练图像和测试图像在此特征脸空间上的投影。接着采用一前向反馈神经网络进行 AU 识别。实验结果如表 7 和表 8 所示。

表 7 单独 AU 的实验结果

AU 号码	平均识别率
1	60.0%
2	50.0%
4	100.0%
5	50.0%
6	50.0%
7	33.3%
9	50.0%
12	40.0%
15	50.0%
17	60.0%
25	75.0%
26	100.0%
27	50.0%

由上表可知,实验中单个 AU 平均识别率为 51.4%,AU 组合平均识别率为 51.4%。而基于事例的 AU 识别算法的平均识别率和相似度分别为 77.5%和 88.2%。因而相对于特征脸方法,实验 1 中,基于事例的 AU 识别方法的结果更为准确。特征脸方法必须先标准化图像,而实验 1 中的图像没有进行任何预处理,并且特征脸方法有图像大小一致、眼睛位置固定等要求,所以对于实时系统中出现的画面,例如实验 2 中使用的图像,就不宜处理。因此不论是针对受控条件下收集的简单图像,还是非受控条件下收集的复杂图像,本文所提出的算法都可以较好识别 AU,具有较好的鲁棒性。

结束语 自动 AU 识别算法研究是表情分析中的一个挑

战性问题,特别是当图像中含有面部附属物等干扰信号时,自动识别更加困难。针对这一问题,本文提出基于事例的 AU 识别算法。该方法无需提取面部特征参数,因而减少了对图像成像环境的要求,具有更好的鲁棒性和实用性。以往研究中,已知 AU 的训练集数据作为本算法的事例集。事例图像的 AU 编码作为解决方案。为了获得一个新的面部图像的 AU 编码,算法采用 IGA,通过交互,充分利用人类的比较能力,快速从事例集中检索一个最接近的事例图像,该图像的 AU 编码即成为新图像的 AU,从而实现 AU 的自动识别。为了验证算法的有效性,本文采用简单面部图像和复杂面部图像进行测试,分别获得 77.5% 和 82.8% 平均识别率,以及 82.8% 和 93.1% 平均相似度,比特征脸方法有较大幅度的提高。

表 8 AU 组合的实验结果

图像号码	AU 组合	平均相似度
1	1+2+5+27	50.0%
2	4+17	100.0%
3	1+2+27	66.7%
4	1+4+15+17	25.0%
5	6+12+25	100.0%
6	4+7+9+17	75.0%
7	1+2+26	66.7%
8	15+17	50.0%
9	1+2+5+26	75.0%
10	12+25	0%
11	4+6+7+9+17	40.0%
12	6+12	0%
13	25	100.0%
14	26	0%
15	6+7+12+25	75.0%
16	12	0%

本文提出的方法不仅为 AU 的自动识别提供了新的解决思路,同时将 IGA 应用到了一个新的研究领域。

虽然文中只对部分 AU 和 AU 组合进行了识别,但是文

中的两个实验具有一定的代表性,可以说明本文方法的有效性、可行性。本文的方法可以推广到任何 AU 和 AU 组合的识别。文中只进行了静态图像的实验,下一步工作是把该方法扩展到动态的图像序列中,并与其他方法进行比较实验。另外,在实验中,我们发现事例集图像的多样性非常重要。如果事例集中没有可匹配的事例,那么识别的效果就会降低。如何交互和动态地增加事例集也是我们下一步的研究目标。

致谢 非常感谢 Dr. Jeffrey Cohn 及其助手为本实验提供了 Cohn-Kanade 数据库,并允许作者使用其中的一些图像作为文章中的图例。其网站是:

<http://vasc.ri.cmu.edu/idb/html/face/facial-expression/index.html>.

参考文献

- Goleman D. Emotion Intelligence. Bantam Books, 1995
- Picard R W. Affective Computing. MIT Press, 1997
- Ekman P, Friesen W. Facial Action Coding System: a Technique for the measurement of facial movement. Palo Alto, CA: consulting Psychologists Press, 1978
- Bartlett M S, Hager J C, Ekman P, et al. Measuring Facial Expression by Computer Image Analysis. Psychophysiology, 1999, 36:253~263
- Donato G, Bartlett M S, Hager J C, et al. Classifying Facial Actions. IEEE trans on Analysis and Machine Intelligence, 1999, 21: 974~989
- Cohn J F, Zlochower A J, Lien J, et al. Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding. Psychophysiology, 1999, 36: 35~43
- Tian Y L, Kanade T, Cohn J F. Recognizing Action Units for Facial Expression Analysis. IEEE Trans on Pattern Analysis and Machine Intelligence, 2001, 23(2): 97~115
- Kapoor A. Automatic Facial Action Analysis. [Master thesis]. MIT, 2002
- Pantic M, Rothkrantz L J M. Facial Action Recognition for Facial Expression Analysis from Static Face Images. IEEE Transactions on Systems, Man, and Cybernetics -Part B, 2004, 34(3): 1449~1461
- Wang Shangfei, Xue jia. Case-based Facial Action Units Recognition Using Interactive Genetic Algorithm. ACII2005, LNCS3784, 2005, 80~87
- Takagi H. Interactive Evolutionary Computation: Fusion of the Capacities of EC Optimization and Human Evaluation. Proceedings of the IEEE, 2001, 89(9): 1275~1296

(上接第 170 页)

步骤 2: 计算核 $\text{Core}(A)$ 。各属性的重要度为: $\text{Sig}_{A-(X)}(X) = \text{Sig}_{A-(S)}(S) = \text{Sig}_{A-(P)}(P) = 9/36$, $\text{Sig}_{A-(M)}(M) = 0$, 于是 A 的核为 $\text{Core}(A) = \{P, S, X\}$ 。 $\text{GK}(\text{Core}(A)) = 12/36$, 故 $\text{Core}(A)$ 为该系统的核心约简, 输出 $\text{Core}(A) = \{P, S, X\}$, 算法终止。

此结果与文[7]利用区分函数得到的约简结果是一致的。不过需要指出的是, 我们给出的算法是尽可能找出不完备信息系统的一个约简, 但一般情况下, 约简不是唯一的, 最小约简也不是唯一的。

结束语 属性约简是粗糙集理论在数据分析上的一个重要应用。一般来讲, 属性约简不惟一。为解决求约简的 NP 问题, 粗糙集研究者们已提出了多种启发式算法。本文针对不完备信息系统, 提出了一种基于知识粒度的属性约简算法, 并分析了该算法的时间复杂度。最后通过例子说明, 使用该算法能找到不完备信息系统的一个约简。

参考文献

- Pawlak Z. Rough sets[J]. International Journal of Computer and Information Sciences, 1982, 11(5): 341~356

- Pawlak Z. Rough Sets-Theoretical Aspects of Reasoning about Data[M]. Dordrecht: Kluwer Academic Publisher, 1991
- Miao Duoqian, Wang Jue. An information-based algorithm for reduction of knowledge[A]. IEEE ICIPS'97[C], 1997. 1155~1158
- 苗夺谦, 胡桂荣. 知识约简的一种启发式算法[J]. 计算机研究与发展, 1999, 36(6): 681~684
- 王国胤, 于洪, 杨大春. 基于条件信息熵的决策表约简[J]. 计算机学报, 2002, 25(7): 759~766
- 梁吉业, 曲开社, 徐宗本. 信息系统的属性约简[J]. 系统工程理论与实践, 2001, 21(12): 76~80
- Kryszkiewicz M. Rough set approach to incomplete information systems[J]. Information Sciences 1998, 112(1-4): 39~49
- Kryszkiewicz M. Rule in incomplete information systems[J]. Information Sciences, 1999, 113(3-4): 271~292
- Liang J Y, Xu Z B. The algorithm on knowledge reduction in incomplete information systems. International Journal of Uncertainty, Fuzziness and Knowledge Based Systems, 2002, 24(1): 95~103
- 何伟, 刘春亚, 赵军, 李华. 不完备信息系统下的属性约简算法[J]. 计算机科学, 2004, 31(2): 117~119
- 黄兵, 周献中, 张蓉蓉. 基于信息量的不完备信息系统属性约简[J]. 系统工程理论与实践, 2005, 25(4): 55~60
- 梁吉业, 李德玉. 信息系统中的不确定性与知识获取[M]. 北京: 科学出版社, 2005. 1~70