

# 计算机智能图像识别算法研究

程咏梅,潘 泉,张洪才,王 刚

(西北工业大学 自动控制系,陕西 西安 710072)

(cymgirl@163.net)

**摘 要** 复杂环境下,要做到对物体进行准确识别,要求计算机视觉系统具有智能化。文中将 D-S 证据推理与不变矩理论相结合,提出了智能图像识别算法,并应用于飞机图像的识别。首先构造了基本置信指派函数,建立了分类规则,然后针对 Dempster 组合规则只适用于高置信度、低冲突性的情况,应用可解决证据冲突的加权分配冲突法及吸收法,完成了计算机飞机图像识别的仿真,比较了三种组合规则的识别效果,结果表明吸收法识别率最高。

**关键词** D-S 证据推理;不变矩;飞机图像的识别

**中图分类号**: TP391;TP393 **文献标识码**: A

## Computer Intelligent Image Recognition Algorithm

CHENG Yong-mei, PAN Quan, ZHANG Hong-cai, WANG Gang

(Department of Automatic Control, Northwestern Polytechnical University, Xi'an Shaanxi 710072, China)

**Abstract:** In complex environment, intelligent computer vision system can meet the needs of accurate recognition object. By combining Dempster-Shafer evidence reasoning with invariant moment theory, this paper presents intelligent image recognition algorithm, and the algorithm is applied to two dimension airplane image recognition. Firstly, a basic belief assignment function is constructed, and a classifying rule is built. Furthermore, in order to solve the problem which Dempster combination rule is used only in case of high belief degree and low conflict, the weighed distribution conflict and the absorptive are used to solve evidence conflict. The result of computer airplane image recognition simulation indicates that absorptive has the highest recognition efficiency in the three combination rules.

**Key words:** Dempster-Shafer evidence reasoning; invariant moment; airplane image recognition

## 1 引言

机器视觉系统在识别图像目标时,一个基本要求是:无论被识别的图像在视场中作何种平移、旋转、放大或缩小,都能正确识别。为解决这个问题,就希望图像特征本身具有不变性。因此,不变性特征的研究一直是感知科学和计算机视觉的研究重点。这方面的研究已取得了一些重要的成果,其中不变矩是一经典的方法<sup>[1-5]</sup>。

不变矩方法就是通过提取具有平移、旋转及比例不变性的数学特征,从而进行图像识别的算法。Hu 首先于 1962 年提出了连续函数矩的定义和关于矩的基本性质<sup>[1]</sup>,在 Hu 研究的基础上,Wong 进一步给出了离散情况下的各阶矩的计算方法<sup>[2]</sup>,图像匹配实验结果表明:比例因子的选取对离散不变矩影响较大。针对这一问题,文献[5]提出了归一化中心矩方法,经过证明一组新的不变矩特征在离散状态下具有平移、旋转及比例不变性。

不确定性推理理论——处理那些具有不完全、不确定、不清晰的信息或数据,是人工智能的一个重要研究领域。关于不确定性推理的成果目前已相当丰富,但 D-S 证据推理<sup>[6-7]</sup>

以其在不确定性的表示、量测和组合方面的优势而受到广泛的重视。经过科学家多年的不断探索,D-S 证据推理理论在不断的得到完善,实际应用也取得了丰硕成果。但是,从国内外目前的研究状况来看,证据冲突的处理方法及衡量标准、证据推理的算法实现、证据推理与其他理论的结合、具体应用背景中基本置信指派函数的构造问题是研究的热点。研究表明产生证据冲突的原因主要是 Dempster 组合规则只适用于高置信度、低冲突性的情况,因此组合规则的改进方法相继提出,其中加权分配冲突法及吸收法在证据高冲突情况下具有良好的性能<sup>[8-9]</sup>。

本文将 D-S 证据理论与不变矩理论相结合,来进行二维飞机图像的识别。证据理论用于这一领域是一个全新的课题。首先针对飞机图像识别这一具体应用背景,计算离散不变矩,构造基本置信指派函数,建立了分类规则,进一步利用 Dempster 组合规则,加权分配冲突法及吸收法完成二维飞机图像的识别。

## 2 Hu 不变矩<sup>[1-2]</sup>

对于二维图像,离散灰度函数  $f(m, n)$ ,相应中心矩的公式如下:

收稿日期:2003-08-14

基金项目:国家自然科学基金资助项目(60372085);陕西省自然科学基金资助项目(2000X17);航空科学基金资助项目(99D53041)

作者简介:程咏梅(1960-),女,陕西西安人,副教授,博士,主要研究方向为:人工智能、图像识别、数字信号处理、复杂环境下机动运动目标建模、数据关联、主动及被动式跟踪。

$$\eta_{pq} = \sum_{m=1}^M \sum_{n=1}^N (m - \bar{x})^p (n - \bar{y})^q f(m, n) \quad (1)$$

其中  $\bar{x}$   $\bar{y}$  是图像  $f(m, n)$  的重心坐标,利用二阶和三阶中心矩构造了 7 个离散不变矩,具体表达式如下:

$$M_1 = \eta_{20} + \eta_{02} \quad (2)$$

$$M_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (3)$$

$$M_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (4)$$

$$M_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (5)$$

$$M_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \\ [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\ [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (6)$$

$$M_6 = (\eta_{20} - \eta_{02}) [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + \\ 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (7)$$

$$M_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - \\ 3(\eta_{21} + \eta_{03})^2] + (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) \\ [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (8)$$

上述 7 个离散不变矩特征不具有比例不变性。为了解决比例因子的影响,文献[5]提出了归一化中心矩方法,经过证明一组新的离散不变矩特征  $M_2^* \sim M_7^*$  具有平移、旋转及比例不变性。它们是:

$$M_2^* = \frac{M_2}{M_1^2} \quad (9) \quad M_5^* = \frac{M_5}{M_1^6} \quad (12)$$

$$M_3^* = \frac{M_3}{M_1^3} \quad (10) \quad M_6^* = \frac{M_6}{M_1^4} \quad (13)$$

$$M_4^* = \frac{M_4}{M_1^3} \quad (11) \quad M_7^* = \frac{M_7}{M_1^6} \quad (14)$$

### 3 D-S 证据推理<sup>[6-7]</sup>

D-S 证据推理是建立在一个非空集合  $\Theta$  上的理论,  $\Theta$  称为辨识框架,  $R$  是辨识框架幂集  $2^\Theta$  中的一个子集,  $A$  为  $\Theta$  的子集,如有集合函数  $m: R \rightarrow [0, 1]$  满足下列条件:

$$\sum \{m(A) | A \subseteq \Theta\} = 1 \quad (15)$$

$$m(\emptyset) = 0 \quad (16)$$

$m$  为辨识框架  $\Theta$  的基本概率指派 BPA( Basic Probability Assignment )函数,也叫基本置信指派 BBA( Basic Belief Assignment )函数,  $m(A)$  表示证据支持命题  $A$  发生的程度,而不包含对  $A$  的真子集的支持。

如果  $A$  为  $\Theta$  的子集,且  $m(A) > 0$ ,则称  $A$  为证据的焦点元素,简称焦点。所有焦点的集合称为核,证据由若干证据体  $(A, m(A))$  组成,利用证据体可以定义  $R$  上的三个测度函数,即置信函数  $Bel$ 、似真函数  $Pl$  和公共函数  $q$ :

$$Bel(A) = \sum \{m(B) | B \subseteq A, B \neq \emptyset\} \quad (17)$$

$$Pl(A) = \sum \{m(B) | A \cap B \neq \emptyset\} \quad (18)$$

$$q(A) = \sum \{m(B) | A \subseteq B \subseteq \Theta\} \quad (19)$$

置信函数  $Bel(A)$  表示给予命题  $A$  的全部支持程度,包括对  $A$  的子集的支持。似真函数  $Pl(A)$  表示不反对命题  $A$  的程度。公共函数  $q(A)$  没有明显的含义,但可以简化运算,如组合

公式的简化。以上三个函数和基本置信指派函数可以利用 Möbius<sup>[6-7]</sup>公式互相导出。

$[Bel(A), Pl(A)]$  构成证据不确定区间,表示证据的不确定程度。证据推理的目的之一就是缩小不确定区间。

两个或多个置信函数可以用 Dempster 组合规则来组合,通过计算基于不同来源置信度的正交和,从而找到一个新的置信函数。假定辨识框架  $\Theta$  上性质不同的两个证据  $A$  和  $B$ ,其焦点分别为  $A_i$  和  $B_j$  ( $i = 1, 2, \dots, n; j = 1, 2, \dots, m$ ),基本置信指派函数分别为  $m_1$  和  $m_2$ ,按照 Dempster 规则组合:

$$m(A) = K^{-1} \{ \sum m_1(A_i)m_2(B_j) | A_i \cap B_j = A \} \\ A \neq \emptyset, A \subseteq \Theta \quad (20)$$

$$m(\emptyset) = 0 \quad (21)$$

其中

$$K = \sum \{m_1(A_i)m_2(B_j) | A_i \cap B_j \neq \emptyset\} = \\ 1 - \sum \{m_1(A_i)m_2(B_j) | A_i \cap B_j = \emptyset\} \quad (22)$$

为归一化因子。由上面的新基本置信指派可以得到新证据体。

当证据之间所支持命题不一致性时,就产生了证据“冲突”问题。证据理论用于数据融合的目的就是使组合后的置信指派函数向辨识框架中更小的子集移动,使某个命题的结论更加明朗。但是它无法处理证据不一致的极限情况。

文献[8-9]在系统分析了已有方法的不足后,认为问题的关键在于组合规则对冲突的处理不当。如式(21)所示,Dempster 规则为了保持基本置信指派函数的归一性,在处理矛盾因子时,使两个证据的公共焦点的基本置信指派变为原来的  $1/K$ ,这意味着把局部的冲突放在全局中去分配,但冲突指派并非所有焦点共同造成的,也并非基本置信指派大的公共焦点产生冲突的可能性大,这样分配会带来不合理性。根据此分析,提出了加权分配冲突法和吸收法。

加权分配冲突法<sup>[8-9]</sup>:假定识别框架  $\Theta$  上性质不同的两个证据  $B$  和  $C$ ,其焦点分别为  $B_i$  和  $C_j$  ( $i = 1, 2, \dots, k; j = 1, 2, \dots, l$ ),对应的基本置信指派函数分别为  $m_1$  和  $m_2$ ,则加权分配冲突法的组合规则为:

$$m(A) = \sum_{B_i \cap C_j = A} m_1(B_i)m_2(C_j) + \\ \sum_{A \cap B_i = \emptyset, A \cap C_j = \emptyset} [w m_1(A)m_2(C_j) + (1-w)m_1(B_i)m_2(A)] \quad (23)$$

比例系数  $w$  可以用两传感器数据可靠程度的比值来定义,也可以用文献[10]提出的鲁棒性分析中鲁棒范围来确定。

修正的吸收法 吸收法<sup>[8-9]</sup>在表述上不够完善,本文对表述形式进行了修正,并没有改变该方法的实质内容。

假定识别框架  $\Theta$  上性质不同的两个证据  $B$  和  $C$ ,其焦点分别为  $B_i$  和  $C_j$  ( $i = 1, 2, \dots, k; j = 1, 2, \dots, l$ ),对应的基本置信指派函数分别为  $m_1$  和  $m_2$ ,则吸收法的组合规则为:

$$m(A) = \sum_{B_i \cap C_j = A} m_1(B_i)m_2(C_j) + \\ \sum_{A \cap C_j = \emptyset, m_1(A) > m_2(C_j) + L} m_1(A)m_2(C_j) + \\ \sum_{B_i \cap A = \emptyset, m_1(B_i) < m_2(A) - L} m_1(B_i)m_2(A) + \\ \sum_{A \cap C_j = \emptyset, m_1(A) - m_2(C_j) \leq L} m_1(A)m_2(C_j)/2 +$$

$$\sum_{B_i \cap A = \emptyset \wedge m(B_i) - m(A) \leq L} m_i(B_i) m_2(A) / 2 \quad (24)$$

其中  $L$  表示指派值的一个限度,当两证据置信指派差在这个限度范围内,认为两证据对冲突的产生起同样的作用,所以冲突置信指派值应均分给两焦点。反之,若置信指派差超过这个范围,则把冲突置信指派分配给具有较大置信指派值的焦点,即大的置信指派吸收了小的置信指派,这就是“吸收法”的含义。

4 飞机图像目标识别

4.1 基本置信指派的构造

选取飞机的类型作为识别框架,即若有  $n$  类飞机  $F_1 \dots, F_n$ , 则  $\Theta = \{ F_1 \dots F_n \}$ 。首先建立这  $n$  类飞机的模型库,即将飞机的某个姿态的图像通过二值化,边缘提取,然后按照式(9)~(14)求取离散矩特征。一组矩特征就对应着一类飞机。当目标图像到来时,同样要二值化,边缘提取后提取矩特征。应用证据理论的困难就在于如何将目标的矩特征和库中各矩特征转换为可以进行推理的基本置信指派函数。

由模式识别的知识可知,两个模式  $x$  与  $y$  相似,就是说  $x$  与  $y$  的特征差不多。设  $x$  与  $y$  所包含的特征有  $m$  个,即

$$\begin{aligned} x &= (x_1 \ x_2 \ \dots \ x_m) \\ y &= (y_1 \ y_2 \ \dots \ y_m) \end{aligned}$$

则定义  $x$  与  $y$  的欧氏距离

$$d = \| x - y \| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_m - y_m)^2} \quad (25)$$

研究表明,从距离函数  $d$  到  $BBA$  函数的映射,即  $R \rightarrow [0, 1]$  映射,是一个非线性映射过程。总的趋势是:与某类飞机的距离越小,分配给该类命题的指派越接近 1;特别地,当距离为 0 时分配给该类命题的指派等于 1;反之,则分配给该类命题的指派越接近 0。因此,需要找到一种函数可以反映这种非线性映射关系,实现从目标识别问题向证据推理问题的转化。指数函数就是具有这种关系的一种函数。设  $r = \exp(-d)$ , 则当  $d \in (\infty, 0]$  时  $r \in [0, 1]$ 。根据这一性质,我们就可以构造基本置信指派了。

根据距离函数,定义证据  $i$  对目标  $F_j$  的基本置信指派:

$$m_i(F_j) = (1 - \alpha) \cdot \exp(-d_{ij}) \quad (26)$$

其中  $\alpha$  为噪声系数。当无噪声时,  $\alpha = 0$ ;当噪声完全淹没了目标时,  $\alpha = 1$ 。

考虑到无噪声时,认为没有不确定性  $m(\Theta) = 0$ ;当噪声完全淹没了目标时,从传感器来的数据无法确认任何目标,即完全不确定  $m(\Theta) = 1$ 。所以构造如下指派:

$$m(\Theta) = \alpha \cdot \prod_{j=1}^N (1 - m_i(F_j)) \quad (27)$$

其中  $N$  表示库中飞机的类型数。

然后对上述基本置信指派进行归一化处理,以满足其定义的要求。

4.2 分类规则

融合后,如何得到识别结果,即对基本置信指派函数的分析与决策没有统一的一般性方法,必须根据具体的问题选择不同的方法。通常以信任度  $\mathcal{J}(A)^{[6-7]}$  作为分类准则。由于本

文的所有命题都是单假设集,所得的置信函数与基本置信指派函数相等,故本文采用基于规则的方法。通过对基本置信指派函数概率意义的分析,确定了若干规则,其中主要的有以下四条:

- 规则 1 :目标类别应具有最大的基本置信指派值;
- 规则 2 :目标类别的基本置信指派值与其他类别的基本置信指派值的差值必须大于某一阈值  $T_1$ ;
- 规则 3 :不确定性置信指派值必须小于某一阈值  $T_2$ ;
- 规则 4 :目标类别的基本置信指派值必须大于不确定性置信指派值。

上述规则 1 表示具有最大信任函数值的类是目标类;规则 2 表示每一证据对所有不同类的支持程度应保持足够大的差异;规则 3 表示对目标类的无知程度或证据的不确定度不能太大;规则 4 表示当对某一目标知之甚少时,不能对其进行分类。

4.3 二维飞机图像识别仿真实验

用 CCD 照相机拍摄飞机图像,在无噪情况下,Adobe Photoshop 绘图软件生成二值图像,大小均存为  $256 \times 256$  点阵。三类不同形状飞机共有 21 幅图像,图 1 至图 7 为第一类,图 8 至图 14 为第二类,图 15 至图 21 为第三类。以第一类为例,第一幅图像为标准图像(即其它图像都是根据其变化而得到的),具体是:第 2、3、4 幅图像分别是第一幅图像放大 1.4、1.8、2.2 倍而得到的;第 5、6 幅图像分别是第一幅图像旋转  $30^\circ$  和  $60^\circ$  而得到的;第 7 幅图像则是第一幅图像平移后得到的。

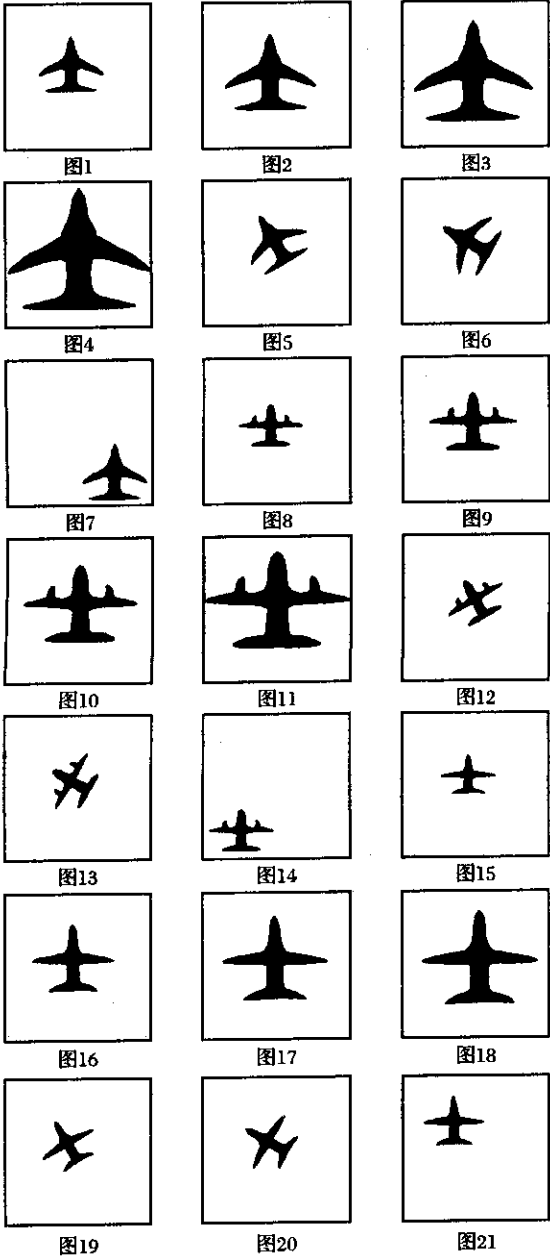
考虑到证据之间的冲突,采用加权分配冲突法和吸收法合成各证据。这两种方法在解决冲突方面显示了较好的性能<sup>[10]</sup>,不仅在低冲突下效果好,高冲突情况下具有同样好的品质。在决策规则中,选取  $T_1 = 0.2, T_2 = 0.5$ 。

表 1 对二维目标的识别结果(修正的 Hu 矩)

图	Dempster 规则	吸收法	加权法(w=0.4)
1	1	1	1
2	1	1	1
3	1	1	0 *
4	0 *	1	1
5	1	1	1
6	1	1	1
7	0 *	1	0 *
8	2	2	2
9	2	2	2
10	2	2	2
11	2	2	2
12	2	2	2
13	2	2	2
14	2	2	2
15	3	3	3
16	3	3	3
17	3	3	3
18	3	3	3
19	3	3	3
20	2 *	2 *	2 *
21	3	3	3
识别率	85.7%	95.2%	85.7%

表 1 给出了采用 Dempster 规则,加权分配冲突法和吸收法合成各证据识别的结果。表中的数字表示识别的目标类型,如“2”表示识别该目标为第二类飞机,数字后的“\*”号表示错误识别,数字“0”表示无法识别。

从识别的结果来看,吸收法的确具有良好的性质,比 Dempster 规则和加权法更有效,而加权法效果的好坏和  $w$  的选取很有关系。



5 结论

从二维飞机图像识别仿真实验可以看出,利用证据推理法对二维飞机图像识别是有效的,特别是吸收法识别效果最好。此处运用的证据推理是一种典型的数据融合方法。数据融合的基本思想就是充分运用传感器提供的冗余信息,从而使得到的结果更为可靠。如果能连续地获得目标不同姿态的图像,用上述的证据推理法能对这些图像进行迭代处理,从而使置信度逐渐增加,错误率逐渐减小。而用传统的最近邻法来处理一系列的图像,需要再增加一个数据融合层,这样速

度上就比直接运用证据推理法慢了许多。因此,基于 D-S 证据推理图像识别算法是智能图像识别算法。

参考文献

[1] Hu MK. Visual Pattern Recognition by Moment Invariant[J]. IRE Trans Information Theory, 1962, 1(8): 179 - 187.

[2] Wong RY. Scene matching with invariant moments[J]. Computer Graphics and Image Processing, 1978, 8(1): 16 - 24.

[3] Dudani SA, et al. Aircraft Identification by Moment Invariants[J]. IEEE Trans., 1977, C - 26(1): 39 - 45.

[4] Pasaltis. Image Normalization by Complex Moments[J]. IEEE Trans. PAMI, 1985, 7(1): 46 - 55.

[5] 潘泉,程咏梅,杜亚娟,等. 离散不变矩算法及其在目标识别中的应用[J]. 电子与信息学报, 2001, 23(1): 30 - 36.

[6] Dempster AP. Upper and lower probabilities induced by a multi-valued mapping[J]. Annals of Mathematical Statistics, 1967, 38(2): 325 - 339.

[7] Shafer G. A Mathematical Theory of Evidence[M]. Princeton University Press, princeton, N. J., 1976.

[8] 张山鹰,潘泉,张洪才. 证据推理冲突问题研究[J]. 航空学报, 2001, 22(4): 369 - 372.

[9] 张山鹰,潘泉,张洪才. 一种新的证据推理组合规则[J]. 控制与决策, 2000, 15(5): 540 - 544.

[10] 潘泉,张山鹰,程咏梅,等. 证据推理的鲁棒性研究[J]. 自动化学报, 2001, 27(4): 338 - 345.

(上接第 55 页)

参考文献

[1] Dumais ST, et al. Using latent semantic analysis to improve information retrieval[A]. CHI'88 Proceedings[C]. 1988. 281 - 285.

[2] Furnas GW, et al. Information retrieval using a singular-value-decomposition model of latent semantic structure[A]. Proceedings of SIGIR'88, 1988. 36 - 40.

[3] Deerwester S, et al. Indexing by latent semantic analysis[J]. Journal of the American Society for Information Science, 1990, 41(5): 391 - 407.

[4] Kolda TG, O'Leary DP. Large latent semantic indexing via a semi-discrete matrix Decomposition[R]. Technical Report No. UMCP-CSD CS-TR-3713, Department of Computer Science, Univ. of Maryland, November 1996.

[5] Li Xiaoli, Shi Zhongzhi. A document categorization based on SOM[A]. PRICAP'2000[C], 2000.

[6] Kohonen T. Self-organization and associate memory[M]. Berlin: Springer-Verlag, 1984, Chapter 5.

[7] Choi D, Park S. Self-creating and organizing neural networks[J]. IEEE Trans. on Neural Networks, 1994, 5(4): 561 - 575.

[8] Dunjia Mladenic. Feature Subset Selection in text-mining[EB/OL]. [http://obelix.ijs.si/Dunjia\\_Mladenic/papers/abstract/apwwEC-MI98.html](http://obelix.ijs.si/Dunjia_Mladenic/papers/abstract/apwwEC-MI98.html), 1998 - 03.

[9] 姜宁,史忠植. 文本聚类中的贝叶斯后验模型选择方法[J]. 计算机研究与发展, 2002, (5): 589 - 593.

[10] 王莉,王正欧. OTGSOM: 一种用于数据聚类的动态自组织神经网络[J]. 电子信息学报, 2003, (4): 678 - 682.

[11] 周水庚,胡运发,等. 隐含语义索引及其在中文文本处理中的应用研究[J]. 小型微型计算机系统, 2001, (2): 987 - 991.