

一种基于数字 AGC 的定点字长控制算法研究与实现

沈智翔, 万明康, 于宏毅

(信息工程大学 信息工程学院, 河南 郑州 450002)

摘要: 对数字信号处理中定点计算字长保留问题进行了研究。首先分析了定点计算中截短量化噪声、保留字长与信噪比损失的关系; 而后针对信号定点处理中的多次字长截短问题, 给出一种基于数字 AGC 的字长截短控制算法, 并在 FPGA 硬件平台上对算法进行了验证。理论分析和验证结果表明, 该算法能够适应输入信号的变化, 保证处理流程中的有效量化位数与量化信噪比。

关键词: 数字信号处理; 定点计算; 有限字长效应; 数字 AGC

中图分类号: TN919

文献标识码: A

文章编号: 1674-7720(2010)11-0010-04

Analysis and implementation of bit-length control algorithm in fix-point processing based on digital AGC

SHEN Zhi Xiang, WAN Ming Kang, YU Hong Yi

(College of Information Engineering, The PLA Information Engineering University, Zhengzhou 450002, China)

Abstract: The retaining of the bit-length in the fix-point calculation of digital signal processing was studied in this paper. Firstly we analyzed the quantization noise caused by multiple bit-length control and the expression of bit-length and SNR loss; then discussed the bit-length control, and gave a bit-length control algorithm based on digital AGC. Finally the algorithm for multiple bit-length control in variable signal processing was proved through the simulation based on FPGA.

Key words: digital signal processing; fix-point processing; finite bit-length; digital AGC

目前以数字信号处理为核心的软件无线电得到越来越广泛的应用。在工程实现中, 数字信号在 FPGA、DSP 以及其他各种数字信号处理芯片中多采用定点计算, 所以处理流程中的定点字长与量化误差是必须考虑的问题。

数字信号采用有限位二进制数表示, 假设 A/D 量化误差为均匀分布白噪声且与信号不相关^[1], 则经 B bit 均匀量化后, 量化信噪比^[2-3]可表示为:

$$SNR_q = 6.02B + 10\lg(12) - 20\lg(P_m/\sigma_x) \quad (1)$$

信号方差为 σ_x^2 , P_m 为量化器的满幅度值。除量化位数外, 量化信噪比同时取决于信号方差与量化器量程的比值, 即信号的有效量化位数。在不溢出的前提下应使信号幅值尽可能接近量化器满量程。

对于零均值高斯信号可取 $P_m/\sigma_x = 3$, 在 8 bit 量化时量化信噪比约为 50 dB。可见在保持一定有效量化位数的情况下量化信噪比是比较高的。但实际应用中信号经

过一级运算后结果往往会超出规定字长, 需要进行截短处理。多次截短产生的误差会不断累积, 同时需要考虑如何保证一定的有效量化位数。如选择简单的取高位截短, 在多级处理后有效量化位数可能远小于字长, 带来信噪比损失和计算效率的降低。因此, 实现时需针对具体算法流程进行定点设计, 这样就大大降低了处理算法的灵活性和通用性。本文通过建立信号多级定点处理模型, 对字长选取和截短问题进行分析, 提出了更为灵活通用的数字信号处理字长选取策略及字长截短控制处理算法。

1 信号模型

连续随机信号经过等间隔采样量化后得到离散随机序列 $x_0(k)$, $k=0, 1, 2, \dots$, 一次信号处理的运算结果可看作该随机序列经过系统后的输出序列: $x_1(k)$, $k=0, 1, 2, \dots$, 输出序列经过字长截短后进入下一级处理。则经 n 次处理后依次得到输出序列 $x_1(k), x_2(k), \dots, x_n(k)$ 。

假设每次截短带来的截短误差相互独立,便可以将截短误差等效为处理流程中依次引入的若干个噪声源,从而得到信号多级定点处理模型如图1所示。

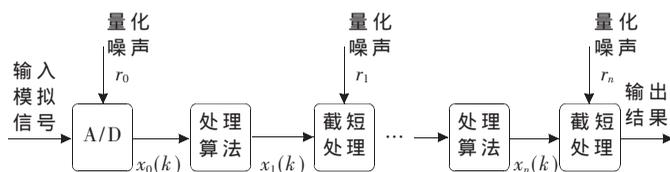


图1 信号多级定点处理流程图

模型中以截短处理为节点,将信号处理流程划分为多级。模型假设A/D满量程采样,采样序列为零均值平稳序列,处理算法均为线性处理。

2 多级处理中的字长选取

当一次运算输出结果超出规定字长时,需要进行尾数截短。截短处理相当于重新量化,可由截短量化噪声的产生、噪声过系统及量化信噪比损失等方面,考虑字长选取问题。

2.1 截短量化噪声分析

式(1)中A/D量化误差取值是连续的,服从均匀分布,而截短量化误差是取值非连续的数字序列。若截去字长为 B_r ,则量化噪声有 2^{B_r} 种取值,可认为其满足离散均匀分布,各取值出现的概率相同,均为 $1/2^{B_r}$ 。

采用舍入法截短,被截短数据为无符号数,最小量化电平为 ΔV ,则量化噪声取值范围为 $(-\frac{\Delta V}{2}, \frac{\Delta V}{2})$ 。

$$\text{均值为: } \sum_{k=-2^{B_r-1}+1}^{2^{B_r-1}} \left(\frac{\Delta V}{2} k\right) \cdot \frac{1}{2^{B_r}} = \frac{\Delta V}{2^{B_r+1}}$$

$$\text{方差为: } \sigma_r^2 = \sum_{k=-2^{B_r-1}+1}^{2^{B_r-1}} \left(\frac{\Delta V}{2} k - \frac{\Delta V}{2^{B_r+1}}\right)^2 \cdot \frac{1}{2^{B_r}} = \frac{(\frac{\Delta V}{2})^2}{2^{B_r}} \cdot \frac{1}{2^{B_r}} \left(\sum_{k=0}^{2^{B_r-1}-1} k^2 + \sum_{k=0}^{2^{B_r-1}-1} k^2 - \sum_{k=-2^{B_r-1}+1}^{2^{B_r-1}} \left(k - \frac{1}{4}\right) \right) = \frac{(\Delta V)^2}{12} \times \frac{(2^{2B_r} - 1)}{2^{2B_r}} \quad (2)$$

若被截短数据为有符号数表示,则对负值数据,量化噪声取值范围为 $[-\Delta V/2, \Delta V/2]$,均值为 $-\Delta V/2^{B_r+1}$ 。根据数据序列的零均值假设,可认为总的量化噪声均值为零,则方差为:

$$\sigma_r^2 = \sum_{k=-2^{B_r-1}+1}^{2^{B_r-1}} \left(\frac{\Delta V}{2} k\right)^2 \cdot \frac{1}{2^{B_r}} = \frac{\Delta V^2}{12} \times \frac{(2^{2B_r} - 1)}{2^{2B_r-1}} \quad (3)$$

由式(2)、式(3)可得,截短量化噪声功率由最小量化电平 $\Delta V = P_m / 2^{B_r}$ 以及截短字长 B_r 决定。随着 B_r 的增加趋近于均匀分布;均值趋近于零,方差趋近于 $\Delta V^2 / 12$ 。

《微型机与应用》2010年第11期

其中 B_r 对量化噪声功率影响较小,截短量化噪声功率主要由截短后的保留字长 B' 决定,对 P_m 归一化,近似取: $\sigma_r^2 = (1/2^{B'}) / 12$ 。

2.2 截短处理的量化信噪比损失

设输入噪声信号为 $r(k)$,均值为 M ,方差为 σ^2 。经过线性系统后的输出噪声序列为 $r_f(k)$,则输出噪声序列均值为: $M_f = M \sum_{m=-\infty}^{+\infty} h(m)$,输出噪声方差为: $E[(r_f(k) - M_f)^2] =$

$\sigma^2 \sum_{m=-\infty}^{+\infty} h^2(m)$ 。根据图1模型,若已知每级处理算法的 $h(k)$,则可以直接计算出输出的累积量化噪声功率。假设经过信号处理流程后信号功率不变,则 N 次截短后的量化信噪比损失等于量化噪声功率的增加。近似取 $\sum_{m=-\infty}^{+\infty} h^2(m) \approx 1$,量化信噪比损失可表示为:

$$SNR_{\text{loss}} = 10 \log \left[\frac{(N(1/2^{B'})^2) / (1/2^B)^2}{12} \right] = 10 \log \left(N 2^{(2B-2B')} + 1 \right) = 10 \log (N 2^{-2\Delta B} + 1) \quad (4)$$

其中 B 为A/D量化位数, B' 为截短后保留字长, $\Delta B = B' - B$ 。不同截短处理次数,不同保留字长下的量化信噪比损失如图2。

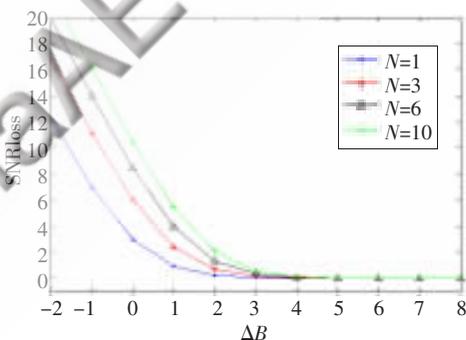


图2 截短后量化信噪比损失

图中 N 为截短次数, N 增加则量化信噪比损失增大;当 ΔB 为负时,量化信噪比损失较大, $\Delta B=2$ 时损失小于3 dB。综上可得,在实际A/D量化信噪比较高的情况下,算法保留字长只要大于A/D量化字长(3~4) bit,就可以大大降低截短量化噪声的影响。

3 数据截短处理算法

假设在 L 位字长中保留 B 位,则保留高 B 位可保证不溢出,但无法保证有效量化数。即量化器的满幅度值与信号均方差的比值 P_m / σ_x 可能变得较大,而信号实际只占用满量程的一部分。

3.1 截短位置的等效增益因子

设输出信号序列直接保留高 B 位得到截短后信号 x ,在不溢出的前提下将截短位置右移 B_a 位则得到截短后信号 $x_{B_a} = 2^{B_a} x$,实际等效为用一个增益因子 $e_x = 2^{B_a}$ 放

欢迎网上投稿 www.pcachina.com 13

大了截短后数据。因此截短位的调整相当于通过该增益因子对截短输出信号幅值进行控制,使得 P_m/σ_x 的取值在一个合理的范围内,以保证有效量化位数。

以零均值高斯序列 $x_0(k)$ 为例,假设 A/D 量化时 $P_m/\sigma_x=3$,自相关函数为 $R_{x_0}(\tau)$,输出序列为 $x_1(k)$,均值 $E[x_1(k)]=0 \times \sum_{m=-\infty}^{+\infty} h(m)=0$,其自相关函数可表示为: $R_{x_1}(\tau)=R_{x_0}(\tau)*h(\tau)*h(-\tau)$,则可取 $e_x=P_m/(3 \times \sqrt{R_{x_1}(0)})$ 。这种计算适用于复杂度较低且 $h(n)$ 确知的系统。而实际中输入信号及处理算法往往并不确定,直接计算较为困难。

3.2 基于数字 AGC 的截短位控制算法

在实际系统中应用的截短处理算法,需要能够应对较复杂的处理流程,适应输入信号的变化。通过建立前馈自动增益控制模型^[4-5]确定增益因子,可较好地完成截短位的自适应调整。首先采用参数估计的方法计算信号均方差;而后将均方差的估计值与参考值 $R=P_m/3$ 比较,其中 α 为误差信号修正系数,通过迭代收敛到最佳增益值。实现流程如图 3。

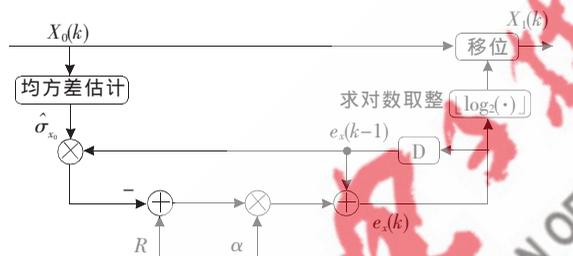


图 3 截短位控制算法流程

模型中增益调整通过移位实现,因此对输出增益值求以 2 为底的对数并取整数部分得到移位长度。根据流程图可得:

$$e_x(k+1)=e_x(k)+\alpha[R-e_x(k)\hat{\sigma}_{x_0}]=e_x(k)[1-\alpha\hat{\sigma}_{x_0}]+\alpha R \quad (5)$$

设信号由零时刻开始输入,可求得 $e_x(k)=R/\hat{\sigma}_{x_0}[1-(1-\alpha\hat{\sigma}_{x_0})^k]$,所以当 $\alpha\hat{\sigma}_{x_0}<2$ 时有 $\lim_{k \rightarrow \infty} e_x(k)=R/\hat{\sigma}_{x_0}$,求对数取整得到移位长度 $B_a=\lfloor \log_2(P_m/3\hat{\sigma}_{x_0}) \rfloor$ 。

若参数估计样本长度为 M ,则 $\hat{\sigma}_x^2=\frac{1}{M} \sum_{k=1}^M X(k)^2$ 是 σ_x^2 的一致估计量,随样本点数增加依概率收敛于真实值,即对于平稳信号可通过增加样本长度将估计偏差控制在允许的范围。计算误差信号时,取参考值 $R=P_m/3$ 适用于高斯信号,实际信号可令参考值略小于 $P_m/3$ 以避免溢出。

实现过程中,复杂度较大仍是该模型的一个主要缺点。尤其在多次截短处理时,为进一步降低计算复杂度,

可以采用直接统计信号动态范围方法代替均方差估计。

设信号均值为零,动态范围表示为 $[+X_m, -X_m]$, X_m 为信号幅度的最大值。以分段统计最大幅值代替均方差估计。若样本长度为 M ,参数估计时需进行 M 次乘累加运算和一次除法运算,而最大幅值统计只要进行 M 次比较。理论参考值可取 $R=2^{B-1}-1$,为保留字长所能表示的最大幅值。与参数估计相比,模型输出随机误差相对较大,需要保证足够的统计长度,并判断剔除错误数据以保证稳健性。

4 FPGA 仿真与实现

下面通过基于 FPGA 的仿真实现在实际信号处理流程中对算法进行验证。经宽带采集的中频信号在 FPGA 中完成数字变频、采样率变换以及数字滤波等处理,基本流程如图 4。

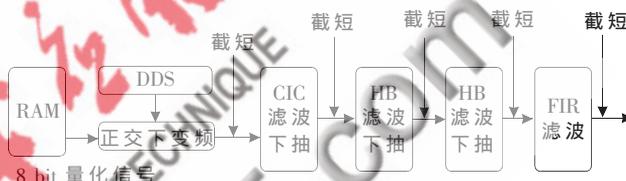


图 4 信号处理定点仿真流程图

仿真信号由载波频率分布于 0.3 MHz~1.4 MHz 范围内的若干路 QPSK 信号组成,码元速率 64 kb/s,升余弦成型,滚降系数 0.5,采样频率 3.072 MHz,8 bit 量化。信号在计算机中生成并量化,以初始化文件的形式下载到 FPGA 的 RAM 中。

FPGA 采用 Altera 公司的 Stratix II EP2S90,读取采样信号后,首先经过正交下变频,数字频率合成模块输出字长为 16 bit。而后进行降采样,包括一级 3 阶 CIC 滤波器和两级 14 阶半带滤波器,将采样率降至 256 kHz,最终经过一级高阶 FIR 低通滤波器提取出一路信号。在正交下变频、降采样和 FIR 滤波器后进行字长截短,每次截短保留 12 bit 字长。所有滤波器基于 FIR Compiler 生成,双通道完成实部与虚部的滤波,输入字长 12 bit,系数字长 16 bit。

正交下变频为 8 bit 与 16 bit 的乘法,输出字长 24 bit;CIC 滤波采用级联的方式增加阻带衰减,输出字长 16 bit;FIR Compiler 在滤波器生成时会根据滤波器系数和输入字长自动确定输出字长;HB 滤波器输出字长 28 bit,FIR 滤波器输出字长 29 bit。则 5 次截短分别需截去 12 bit、4 bit、16 bit、16 bit、17 bit。算法中求对数取整运算通过查找表实现,在 SignalTap II Logic Analyzer 中采集仿真结果。表 1 给出 3 组不同信号收敛后的移位长度 B_a 。

信号一为功率相等的 2 路 QPSK 信号,中心频率分别为 1 MHz 和 1.1 MHz,下变频频点为 1 MHz;信号二为 8 路功率相等的 QPSK 信号,中心频率分布于 0.3 MHz~

表 1 3 组信号截短方案仿真结果

信号	B_n /截短位置	DDC	CIC	HB1	HB2	FIR
一	参数估计	1	1	0	0	1
	最大值统计	1	1	0	0	1
	参考值	1	1	0	0	1
二	参数估计	1	2	1	1	1
	最大值统计	1	2	1	1	1
	参考值	1	2	1	1	1
三	参数估计	2	2	0	2	2
	最大值统计	2	2	0	2	2
	参考值	2	2	0	2	2

1.4 MHz 范围内,下变频频点为 960 kHz;信号三在信号二的基础上随机调整了信号功率。误差信号修正系数取 $\alpha=0.5$,参数估计模型中采用滑动窗进行方差估计,滑动窗长取 500 点,最大值统计模型中分段统计长度为 500 点。2 种模型取整后的输出结果与参考值一致,参考值为由信号实际取值范围得到的理想截短方案,可以看出在不同的信号条件下,理想移位长度 B_n 是不同的。若进行取高位截短,对于信号一将损失 3 bit 有效量化位,而对于信号三将损失 8 bit 有效量化位。

为进一步检验模型的收敛跟踪性能和输出结果的方差,将取整前的输出结果进行对比。采集查找表前的输出数据,在 Matlab 中取对数并做出曲线图。2 种模型样本长度 M 取 500 点和 200 点时的跟踪收敛曲线如图 5。

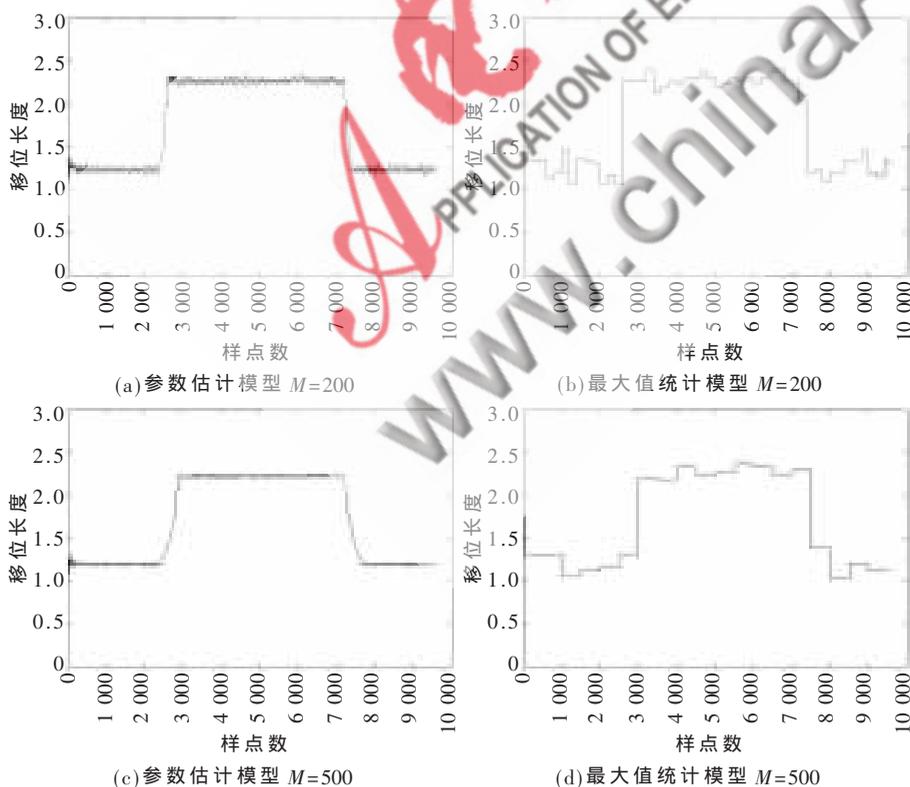


图 5 字长控制算法取整前输出

图 5 中横坐标为样点数,纵坐标为 DDC 后截短移

位长度(取整前)。输入信号在 2 400 采样点处由信号二变为信号三,在 7 200 点处由信号三变回信号二。样本长度相同时,参数估计模型抖动较小,而动态范围统计模型则相对较大,相比较而言前者更加充分地利用了样本数据,受随机性影响较小。参数估计模型收敛跟踪速度与滑动窗长有关;最大值统计对于信号动态范围增大的情况能够快速跟踪从而避免溢出(7 200 点处),而信号动态范围减小时输出有一个样本长度左右的延迟(2 400 点处)。

本文从两个方面分析了定点计算中的有限字长保留问题。首先将字长的截短等效为重新量化,分析了量化信噪比损失及计算中所需的最小保留字长;而后讨论了最佳截短方案及其实现,将截短位选择等效为信号增益调整,利用自动增益控制模型完成实时信号截短处理;最后通过硬件仿真,在典型软件无线电处理流程中对算法进行了验证,并对比讨论了基于参数估计与最大值统计两种模型的计算量和算法性能特点。本文关于定点计算有限字长保留问题的分析和实现,在工程实现中有实际参考价值和应用价值。

参考文献

- [1] WIDROW B, KOLLAR I, LIU M C. Statistical theory of quantization [J]. IEEE Transactions on Communications, 1964, 12(4):162-165.
- [2] 陈亮,杨吉斌,张雄伟.信号处理算法的实时 DSP 实现[M].北京:电子工业出版社,2008.
- [3] 丁玉美.数字信号处理[M].西安:西安电子科技大学出版社,2000:262-290.
- [4] GUESALAGA A, TEPPER S. Synthesis of automatic gain controllers for conical scan tracking radar [J]. IEEE Transactions on Aerospace and Electronic Systems, 2000, 36 (1):302-309.
- [5] JOHNSON C R. 软件无线电[M].北京:机械工业出版社,2008.

(收稿日期:2010-01-07)

作者简介:

沈智翔,男,1985 年生,硕士研究生,主要研究方向:无线通信。

万明康,男,1983 年生,博士研究生,主要研究方向:无线通信。

于宏毅,男,1963 年生,博士生导师,教授,主要研究方向:无线通信、通信信号处理。

基于本体的概念相似度计算及其应用*

冉 婕^{1,2}, 孙 瑜¹, 漆丽娟²

(1. 云南师范大学 计算机科学与技术学院, 云南 昆明 650092;

2. 云南昭通师范高等专科学校 计算机科学系, 云南 昭通 657000)

摘要: 提出了基于语义相似度和相关度的综合概念相似度计算方法。语义相似度考虑了语义距离和本体库特征, 加入概念的信息量、概念的深度、概念的密度和不对称因子的辅助影响; 语义相关度从直接相关、间接相关、直接继承和间接继承几个方面考虑。通过实验和两种传统的语义相似度计算方法进行对比, 本方法能更好地地区分本体树中不同关系的概念对, 验证了该方法的有效性。

关键词: 本体; 相似度; 语义距离; 信息量

中图分类号: TP391

文献标识码: A

文章编号: 1674-7720(2010)11-0014-03

Concept similarity computation based on ontology and its application

RAN Jie^{1,2}, SUN Yu¹, QI Li Juan²

(1. Institute of Computer Science and Information Technology, Yunnan Normal University, Kunming 650092, China;

2. Department of Computer Science, Zhaotong Teacher's College, Zhaotong 657000, China)

Abstract: This paper puts forward an integrated method based on semantic similarity and semantic relevancy. In semantic similarity, we think about the semantic distance and the characteristics of ontology, the concept's amount of information, the depth of concept, the density of concept and symmetry factor; in semantic relevancy, we think about the directly relation and indirectly relation, direct inheritance and indirect inheritance. Compared with two traditional semantic similarity computation methods, this method can better distinguish the different concept in ontology tree and this method is effective.

Key words: ontology; similarity; semantic distance; amount of information

目前, 信息检索大多基于关键字进行, 查准率及查全率均不高, 而本体能描述数据的语义, 基于本体进行信息检索, 检索效率显然要高。参考文献[1]指出, 本体在信息检索中的应用能够显著地提高检索的精确率和返回率。在信息检索领域中, 概念的语义相似度计算起着重要的作用, 因此可以利用本体计算概念间的语义相似度。语义相似度在不同的应用领域中可能会有不同的含义。在信息整合领域中, 相似度一般指的是文本与文本能够匹配的程度; 而在信息检索领域中, 相似度则反映与用户查询在语义上的匹配程度。相似度越高, 表明该文本与用户的请求越接近^[2]。本文的研究背景为基于本体的信息检索。

利用本体计算概念间相似度的基础是: 2 个概念

间具有一定的语义相关性, 它们在概念间的结构层次网络图中存在一条路径^[3]。Resnik^[4]根据 2 个词的公共祖先节点的最大信息量来衡量 2 个词的语义相似度; Agirre^[5]在利用 WordNet 计算词语的相似度时, 考虑了语义距离、概念层次树的深度和概念层次树的区域密度; 参考文献[6]提出基于距离的语义相似度计算模型, 这种模型简单直观, 但它依赖于预先建立好的本体层次网络; 参考文献[7]引入计算语言学中的语义距离思想来计算概念相似度, 但其考虑概念间的相似度影响因素较少。针对上述研究情况, 本文提出了一种基于语义相似度及相关度的综合概念相似度计算方法。

1 概念相似度计算

当 2 个概念具有某些共同特征时, 则定义它们是相似的, 用 $\text{sim}(x, y)$ 表示概念 x, y 之间的相似度。形式上, 相似度计算满足^[7]: (1) 相似度的值为 $[0, 1]$ 区间中的一

* 基金项目: 国家自然科学基金项目(60903131); 云南省社会发展科技计划应用基础研究项目(2009ZC052M); 云南省教育厅重点项目(07Z10661)