

# 基于语料库的汉语复合名词短语自动获取研究

王 萌<sup>1</sup>, 朱 虹<sup>2</sup>, 徐 戈<sup>3</sup>

(1.江南大学 人文学院教育技术系,江苏 无锡 214122;

2.中国标准化研究院,北京 100191;3.闽江学院 计算机科学系,福建 福州 350108)

**摘 要** 汉语复合名词短语广泛存在于各种文体中,且绝大部分是低频的,这给复合名词短语的自动获取带来了很大的挑战性。文章针对统计指标不能有效获取低频复合名词短语的问题,提出了新的解决方法,将其视作一个分类问题,利用统计指标获取典型的、高频的复合名词短语作为训练数据,抽取多种特征,来帮助发现低频的复合名词短语,实验结果说明该思路是有效的。

**关键词** 复合名词短语;自动获取;SVM;统计指标

中图分类号:TP391 文献标志码:A 文章编号:1009-8666(2014)12-0057-06

## A Study of Automatic Acquisition of Chinese Compound Noun Phrases Based on Corpus

WANG Meng, ZHU Hong, XU Ge

(Department of Educational Technology, School of Humanity, Jiangnan University, Wuxi Jiangsu 214122; China National Institute of Standardization, Beijing 100191; Department of Computer Science, Minjiang University, Fuzhou Fujian 350108, China)

**Abstract:** Chinese compound noun phrases are very common in most texts. Automatic acquisition of these compounds are extremely difficult because most of them have relatively low frequency in corpus. This paper mainly aims at the problem that statistical indexes can not be effective for low frequency complex noun phrases and proposes a new method to deal with it. In this new method, the problem is modeled as a classification problem, which makes use of the statistical indexes to get the typical, frequent compound noun phrases as training data. Then the extraction of various features help find the infrequent ones. The experimental results show that this proposed idea is applicable.

**Key Words:** Noun Compounds; Automatic Acquisition; SVM; Statistical Indexes

### 1 复合名词短语概述

复合名词短语(noun compounds)是一种特定类型的短语,它由相邻的名词序列组成,其功能整体上相当于一个名词<sup>[1]</sup>,如“电子警察”“啤酒公司”

“水果价格”和“空气质量问题”。通常把复合名词短语中的最后一个名词称为中心词(head),前面的成分称为修饰词(modifier)。从语法角度来说,复合名词短语和词比较相似,整个复合名词短语的功能相当于中心名词的功能。

收稿日期:2014-09-25

作者简介:王萌(1977—),女,河南洛阳人。江南大学讲师,博士,研究方向:计算语言学;朱虹(1981—),女,江苏苏州人。中国标准化研究院助理研究员,研究方向:信息处理、数据标准化。

复合名词短语广泛存在于英语、汉语、日语、德语等各种语言中,主要有以下几个特点:

第一,两个词语通过“复合”这一操作来实现信息压缩,并以最少的文本呈现,这符合语言学中的“经济原则”(人们在交际时总是力求用最简洁的语言来表达复杂完整的意思),因此复合名词短语的使用非常普遍,经常出现在如新闻、科技报告以及小说等各种文体中。

第二,复合名词短语有很强的能产性。一方面,在日常生活中,人们根据交流的需要,不断地产生新的复合名词短语,如“傻瓜相机”“板凳队员”“凤凰男”等,来指称一类新出现的事物。另一方面,在快速更新的科技领域中,专业术语通常以复合名词短语的形式出现。

第三,虽然复合名词遵循非常简单组合规则(相邻的名词序列),但是它们在句法分析和语义解释上表现出较高的歧义性。特别是在汉语中,复合名词短语的分析效果对整个句法分析有着较大的影响。

这些特点使得复合名词短语在语言学和计算语言学领域成为一个热点研究课题,其涉及到的研究范围也越来越广泛,包括复合名词短语的自动获取、句法分析、语义解释、翻译以及语义焦点分析等等。本文主要关注复合名词短语的自动获取研究。

## 2 问题提出

早期的研究者通常将复合名词短语的获取问题归入术语或搭配的认识,但是两者有着本质的不同。根据 Downing 的定义,复合名词短语只能包含名词<sup>[2]</sup>,而术语或者搭配包含的词类更广泛,如形容词、动词和介词等。此外,复合名词短语是连续的词语序列,而搭配可以是长距离的、不连续的。获取复合名词短语的方法通常有两种,即规则和统计的方法。在规则的方法中,研究者利用表层的句法信息来制定一些规则<sup>[3-4]</sup>,如词性或词语形态信息,去判断一个词语序列是否构成复合名词短语。在统计的方法中,研究者用一些统计指标(如互信息和对数似然比)来衡量词语序列构成复合名词短语的可能性<sup>[5-7]</sup>。

复合名词的衍生能力很强,根据人们交流的需要,新的复合名词几乎每天都在产生,新产生的复合名词给自动获取任务带来了很大的挑战性,尤其是对于常用的统计方法而言。众所周知,统计

方法中一个基本假设是两个有关联的词语共现的次数比偶然共现的次数要多,这就要求复合名词短语必须在语料中有一定的出现次数。否则,使用各种统计方法都不能得到可信的结果。但是,这几乎是一对矛盾,复合名词短语中绝大部分是低频的,而统计的方法对低频的数据不能给出准确的预测结果。因此,本文主要研究如何获取低频的复合名词短语,将该问题视作一个二分类问题,即判断“名-名”序列是不是复合名词短语。首先,用三种统计量获取高频的复合名词短语,将其作为训练语料,抽取多种特征;其次,用 Support Vector Model(SVM)分类器进行学习,用于预测低频的“名-名”序列是否能构成复合名词短语,实验结果说明该思路是有效的。

## 3 数据准备

首先,从语料中抽取“名-名”序列作为复合名词短语的候选,实验的语料是 1998 年上半年基本标注的《人民日报》。抽取过程包括以下两步:

1.寻找两个邻接的名词序列,且该序列前后相邻的词语不能再是名词。

2.过滤掉包含字母和数词等非汉字字符的名词序列,如“PC 机/n 原理/n”和“X 光/n 机器/n”等。

按照上述步骤,共得到总数为 127 701 词次,包含 48 283 词型的候选二元复合名词短语。表 1 给出了二元候选复合名词的分布情况,绝大部分(大于 71%)的候选出现只出现 1 次,频次大于 5 的仅占到 10%,这样的分布符合齐夫定律(Zipf's Law)<sup>[8]</sup>。

表 1 二元候选复合名词短语分布

出现频次	>=5	>=2	>=1	=1
词型个数	4 140	13 886	48 283	34 397

由此可见,如果我们想尽量多地获取复合名词短语,如何有效地处理这些低频的复合名词短语是一个不可避免的问题。

## 4 三种统计指标

统计指标,如互信息和对数似然比(Log likelihood ratio)等,已经被广泛地用于术语或搭配的获取。本节将探讨三种统计指标:共现频次(CoocF, Co-occurrence Frequency)、互信息(MI, Mutual Information)和依赖率(DR, Dependent Ratio),分析它们在获取复合名词短语方面的表现。

#### 4.1 共现频次

共现频次的基本假设是：如果两个词经常在一起出现,意味着它们联系较为紧密。术语获取的早期研究已经指出共现频次能够很好地体现术语性(termhood)<sup>[4]</sup>。Lapata 对英文复合名词的获取进行了详细研究,实验证明共现频次的表现优于互信息等指标<sup>[9]</sup>。表 2 给出了共现频次最高的部分候选复合名词,其中大部分都是正确的复合名词<sup>①</sup>。

表 2 共现频次最高的候选复合名词短语

候选复合名词短语	CoocF ↓	MI	DR
领导干部	1106	5.758 3	0.286 7
人民群众	911	4.661 3	0.157 2
国际社会	544	4.003 7	0.096 3
* 国关系	510	4.606 4	0.122 6
金融机构	415	5.114 4	0.193 9
* 国人民	381	3.896 9	0.109 8
金融危机	362	5.437 4	0.311 0
记者招待会	360	5.896 9	0.212 4
商业银行	262	5.790 7	0.215 1
基础设施	225	5.592 1	0.376 3

#### 4.2 互信息

作为衡量词语相关性的指标,互信息被广泛地用于搭配抽取。互信息的定义见公式(1),它反应了同时观察到  $n_1$  和  $n_2$  的概率(联合概率)与独立观察到它们的概率的比值。互信息能够用来表示一个事件中蕴含的关于另一个事件的信息量。

$$I(n_1, n_2) = \log_2 \frac{P(n_1, n_2)}{P(n_1) \times P(n_2)} \quad (1)$$

其中  $P$  表示概率,一般用最大似然估计(MLE, Maximum Likelihood Estimation)来计算  $P(n_1, n_2)$ ,  $P(n_1)$  和  $P(n_2)$ , 见公式(2)及(3):

$$P(n_1, n_2) = \frac{freq(n_1, n_2)}{N} \quad (2)$$

$$P(n_i) = \frac{freq(n_i)}{N} \quad (i=1, 2) \quad (3)$$

$req(n_1, n_2)$  表示  $n_1$  和  $n_2$  的共现频次,  $freq(n_i)$  是  $n_i$  出现次数,  $N$  是语料的大小(总词次)。较大的  $I(n_1, n_2)$  值说明两个词语相关度高,反之,它们可能是偶然出现在一起。表 3 给出了互信息最高的部分候选复合名词短语。

#### 4.3 依赖率

根据复合名词短语能产性高的特点,本文提出一个新的统计量,用来衡量短语内组成成分之间的依赖程度,称之为依赖率。对于一个候选复合名词  $n_1 n_2$ , 依赖率定义如公式(4)所示:

表 3 互信息值最高的候选复合名词短语

候选复合名词短语	MI ↓	CoocF	DR
脊髓灰质炎	13.064 6	7	0.866 0
* 厂规厂纪	12.581 9	9	0.666 7
鲨鱼软骨素	12.486 6	8	0.741 6
算术平均数	12.399 6	8	0.516 4
交响诗大合唱	12.176 5	6	0.400 0
民法通则	12.127 7	16	0.677 3
* 氨氮	12.079 1	6	0.444 5
柳木拐杖	12.074 2	13	0.580 5
* 赃款赃物	11.945 3	8	0.514 3
驱逐舰舰长	11.929 6	7	0.436 4

$$DR(n_1 n_2) = \sqrt{LD(n_2) \times RD(n_1)} \quad (4)$$

其中,  $LD$  和  $RD$  定义见公式(5)和(6):

$$LD(X) = \frac{\max_{w \in Left(X)} freq(wX)}{freq(X)} \quad (5)$$

$$RD(X) = \frac{\max_{w \in Right(X)} freq(Xw)}{freq(X)} \quad (6)$$

其中  $freq(X)$  是词语  $X$  在语料中的出现频次,  $Left(X)$  和  $Right(X)$  分别表示词语  $X$  的所有左邻接词和右邻接词的集合,  $w$  表示集合中的一个元素。公式  $LD(X)$  体现了  $X$  左邻接词的多样性, 值越高, 说明其左邻接词变化越小, 相对固定。反之,  $LD(X)$  值越低, 说明  $X$  的左邻接词种类越多, 变化越大。对于词语  $X$  的  $RD(X)$  值, 可以按照同样的方法进行解释。

前文曾提到过“复合”是一个衍生能力很强的语言现象。正如汉语中有一些能产性很高的名词作为中心词, 与其它名词组合形成复合名词短语, 如“先生、公司、系统”等。以“系统”为例, 发现语料中有“电脑系统、公安系统、交通系统”等各种各样的系统, 这种现象会使得  $LD(\text{公司})$  值相对较低。因此, 可以利用词语  $X$  的  $LD(X)$  值来衡量其作为中心词的能产性。此外, 汉语中也存在一些非常活跃的修饰词, 如“国际、经济、科技”等, 它们倾向于充当其它名词的限定成分, 如“国际社会、国际标准、国际会议”等, 通常这类名词的  $RD$  值相对较低。因此,  $RD(X)$  值可以用来衡量词语  $X$  充当修饰语的能力。

将以上两个因素综合考虑, 就得到了最终的依赖率公式(4), 该公式可以用来发现由能产性较高的中心词或者修饰语组成的复合名词短语。表 4 给出了依赖率最小的部分复合名词短语的样例。可以看出, 表中有一部分共现频次小于 5 的复

合名词短语,值得注意的是,对于这部分低频复合名词短语,前两个指标(共现频次和互信息)表现欠佳。而依赖率这一指标对于频次并不敏感,可以有效地获取低频和高频的复合名词短语。

表4 依赖率最低的候选复合名词

候选复合名词短语	DR ↓	CoocF	MI
议长先生	0.032 0	4	4.181 4
全省系统	0.034 8	1	1.239 2
记者先生	0.038 5	1	-1.081 2
新华书店系统	0.038 7	1	3.848 3
总统先生	0.039 0	12	2.777 2
计算机系统	0.039 2	15	4.384 5
法院系统	0.040 9	5	3.217 0
全县工业	0.041 1	1	1.733 1
供销社系统	0.041 3	1	3.981 9
市长先生	0.041 7	4	3.483 1

## 5 基于机器学习模型 SVM 的实验

### 5.1 特征选取

实验共用到三类特征:数值特征、语义特征和字特征,下面分别予以说明。

#### 数值特征(Numerical Feature)

数值特征一共包含4种从训练样本中获取的数值,因为每个统计指标的准确率都在90%以上,所以由它们获取的训练样本能够提供可靠的数值信息。4类特征分别是:

- $f(n, M)$ , 名词  $n$  出现在修饰词( Modifier)的次数
- $f(n, H)$ , 名词  $n$  出现在中心词( Head)的次数
- $P(n, M)$ , 名词  $n$  作为修饰词概率
- $P(n, H)$ , 名词  $n$  作为中心词的概率

其中  $P(n, M)$  和  $P(n, H)$  的定义如公式(7)和(8)所示。

$$P(n, N) = \frac{f(n, M)}{f(n)} \quad (7) \quad P(n, H) = \frac{f(n, H)}{f(n)} \quad (8)$$

其中  $f(n)$  是名词  $n$  在训练样本中的出现次数,与其出现位置(修饰词或中心词)无关。这些数值特征体现了一个名词作为修饰词或中心词的可能性,较高的  $P(n, M)$  说明名词  $n$  经常作为修饰语。但是,数值特征只提供了表层的名词作为修饰语或者中心语的统计信息,它们并没有反映词语在进行“复合”时语义组合的规律。一个解决的方法就是根据某种语义分类体系引入词语的语义

类,让学习算法获取语义组合的规律。

#### 语义特征(Semantic Feature)

把两个词语“复合”在一起的目的是将两者关联起来,使其更容易与同类事物区别,因此,应该有某种语义联系参与其中。例如,“木头桌子、金属盆子、塑料瓶子”等复合名词短语,修饰词和中心词之间都隐含了“制造”这种关联关系。可以将这些复合名词短语视作“材料”和“制造物”进行语义组合而产生的实例,其中修饰词表示“材料”,中心词表示“制造物”。

本文用《中文语义词典》(Chinese Semantic Dictionary, CSD)提供语义信息,词典名词库中共收录37 675个词语,分别描述了它们的语义层级和搭配等信息<sup>[10]</sup>。在CSD中,平均一个名词有1.07个语义类。例如,名词“木头、金属”属于“材料”类,而“桌子、盆子”属于“工具”类。

一个名词可能属于多个语义类,本文在使用语义类信息时并没有对词语进行词义消歧。对于这个问题,本文采取一个较为简单的处理办法:对于每一个多义名词,将其所有的语义类进行合并,形成一个新类。例如,名词“梨”有两个语义类:植物和食物,将两者合并形成“植物+食物”类作为“梨”的语义类。固然,这样的处理方式会丢失一些语义信息,因为它没有反应合并得到的新类与原始语义类之间的重合情况,而是作为两个独立的特征。但是,通过统计,多义名词仅占整个名词集合的7%。

模型使用了一元和二元的语义特征,一元特征是每个词语的语义类,二元特征是两个名词语义类的组合。如果名词没有出现在CSD中,用“NULL”作为其语义特征。

#### 字特征(Character Feature)

字特征也是非常重要的一类特征,包括两个方面:一个是名词本身,一个源于构词法(word formation)。汉语中指示同一类事物的名词大部分都是定中结构,具有相同的核心成分。本文将每个名词的最后一个字视作核心成分,将核心成分作为特征。这其实也是一种回退的方法,可以缓解数据稀疏问题。

模型使用一元特征和二元特征,一元特征是每个词语本身及其核心成分,二元特征包括两个词语的组合及其核心成分的组合。

### 5.2 实验

整个实验的流程如图1所示,用三个统计指



标获取的复合名词作为训练数据，然后随机挑选了 484 个频次为 1 的“名-名”组合作为测试。实验使用的是 LIBSVM 工具包<sup>②</sup>，采用线性核函数和其它默认参数设置。

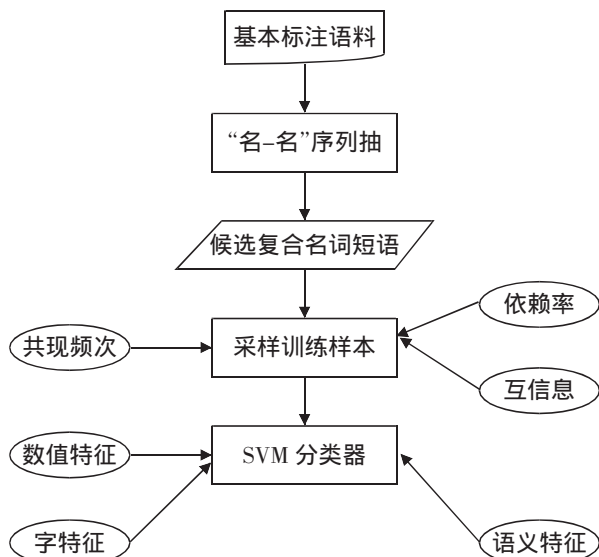


图 1 复合名词短语获取流程

### 5.3 实验结果及分析

本文在单独使用及组合使用各个特征的条件 下进行了实验，实验结果的准确率见表 5。为了比 较，将所有测试数据都标记为正确的复合名词时的 准确率作为基准(baseline)。可以看出，baseline 相当高(78%)，这意味着按照第三节中介绍的抽 取方法，汉语中“名-名”序列的组合很可能是复合 名词短语。这和英文的情况大不相同，Lapata 用同 样的策略得到的英文的 baseline 只有 55%<sup>[9]</sup>。这种 差别显示了汉语以更灵活地方式产生复合名词。

实验结果显示，同时使用字特征和语义特征 时，系统得到最好的结果，而数字特征的表现很不 稳定。汉语是一种意合语言(meaning-combination language)，词语在复合名词中的出现位置非常灵 活。例如，前文提到过“公司”是一个衍生能力很强 的中心词，组成如“啤酒公司、电脑公司、食品公

表 5 各种特征及其组合的准确率

特征	准确率 (%)
Baseline	78.09
Numerical	78.09
Character	80.99
Numerical + Character	78.51
Character+ Semantic	81.61
Numerical + Character + Semantic	77.06

司”等复合名词短语，同时它也经常出现在修饰语 的位置，如“公司职员、公司经理、公司制度”等。因 此，这就造成数值特征的区分能力较弱。当三种特 征同时使用时，性能最差，甚至低于 baseline，这时 数值特征成为噪音，影响整个系统性能。

## 6 结语

本文对复合名词短语的获取问题做了初步的 研究和分析，主要关注的是低频复合名词短语的 获取。我们将其视为一个分类问题，探索了三类特 征的表现，实验结果表明两个名词在组成复合名 词短语时，它们之间的语义联系起了关键作用。对 于高频复合名词短语，很多统计指标都可以有效 地获取，对于低频复合名词短语，在语料中可以观 察到的语言现象较少，各种统计指标表现欠佳。因 此，本文的一个基本思路就是：利用高频复合名 词短语来帮助发现低频的复合名词短语。实验结果 表明，该思路是有效的。

实际上，复合名词的获取问题依赖于很多其 他因素，如世界知识、说话者意图、交流目的以及 文化背景等。然而，将以上这些信息形式化地表示 为特征并非易事。本文将该问题视为静态问题，所 以模型中并没有使用“名-名”序列出现的上下文 信息。下一步的工作，我们将充分发掘这些上下文 环境以获取更多的复合名词短语。此外，还可以尝 试其它的统计方法，得到更多的训练样本，研究该 方法的学习曲线。

注释：

①有 \* 标记表示是错误的复合名词短语，以下类同。

②工具包下载网址 <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>。

参考文献：

- [1]Downing, Pamela. On the Creation and Use of English Compound Nouns [J]. Language, 1977, 53(4): 810-842.
- [2]Downing, Pamela. On the Creation and Use of English Compound Nouns [J].Language, 1977, 53(4): 810-842.
- [3]Lauer, Mark. Corpus statistics meet the noun compound: some empirical results [C].33rd annual meeting of the Association for Computational Linguistics, 1995.
- [4]Justeson, John S and Slava M Katz. Technical Terminology: Some Linguistic Properties and an Algorithm for Identification in Text [J]. Natural Language Engineering, 1995 (1): 9-27.
- [5]Manning, Christopher D., and HinrichSchutze. Foundations of statistical natural language processing [M].Cambridge, MA: MIT Press, 1999.

- [6]Church ,K. and Hanks P. Word Association Norms ,Mutual Information and lexicography [C]//Proceedings of the 27th Annual Meeting of the Association for computational Linguistics ,Vancouver ,Canada ,1989.
- [7]Dunning ,Ted E. Accurate Methods for the Statistics of Surprise and Coincidence [J]. Computational Linguistics ,1993 ,19(1) 61 - 74.
- [8]Zipf ,George Kingsley. Selected Studies of the Principle of Relative Frequency in Language [M]. Cambridge (Mass.) ,1932.
- [9]Lapata ,Maria. The Acquisition and Modeling of Lexical Knowledge :A Corpus-based Investigation of Systematic Polysemy [D]. PhD Thesis. University of Edinburgh 2001.
- [10]王惠 ,詹卫东 ,俞士汶. 现代汉语语义词典规格说明书[J]. 汉语语言与计算学报 ,2003 ,13(2) :159 - 176.
- [11]Chomsky ,N. and Halle ,M. The Sound Pattern of English.[M] 2nd Edition. MIT Press ,Cambridge ,MA ,1991.
- [12]Fazly ,Afsaneh and Paul Cook and Suzanne Stevenson.Unsupervised type and token identification of idiomatic expressions [J].Computational Linguistics.2009 ,35 (1) :61-103.
- [13]Lapata ,Maria and Alex Lascarides. Detecting Novel Compounds :The Role of Distributional Evidence [C]//Proceedings of the 11th Conference of the European Chapter for the Association of Computational Linguistics 2003.
- [14]Lauer ,Mark.Designing Statistical Language Learners :Experiments on Compound Nouns [D].Ph.D.thesis ,Macquarie University. Australia ,1996.
- [15]Levin ,Judith. The Syntax and Semantics of Complex Nominals [M].Academic Press ,New York ,1978.
- [16]Su ,Keh-Yih ,Ming-Wen Wu and Jing-Shin Chang.A corpus-based approach to automatic compound extraction [C]//Proceedings of the 32nd annual meeting on Association for Computational Linguistics ,Las Cruces ,New Mexico ,1994.
- [17]俞士汶 ,段慧明 ,朱学锋 ,等. 北大语料库加工规范 :切分词性标注注音[J]. 汉语语言与计算学报 ,2003 ,13(2) :121-158.
- [18]Meng Wang ,Chu-Ren Huang ,Shiwen Yu ,Weiwei Sun. Automatic Acquisition of Chinese Novel Noun Compounds [C]. Proceedings of The International Conference on Language Resources and Evaluation (LREC). Malta 2010.
- [19]Chen Yirong ,Qin Lu ,Wenjie Li ,Zhifang Sui ,Luning Ji. A Study on Terminology Extraction based on Classified Corpora[C]//Proceedings of the International Conference on Language Resources and Evaluation ,Genoa. Italy 2006.

[责任编辑:王兴全]

---

(上接第 56 页)

便会增多。野生与栽培重楼中元素含量的不同,可能与其生长土壤中元素含量高低有关<sup>[20]</sup>。因此在重楼的种植过程中要根据重楼的需肥特性及生长

规律有针对性地施肥,从而使重楼根茎的品质达到最优。

#### 参考文献:

- [1]胡光万 ,雷立公. 出自深山的良药—重楼[J]. 植物杂志 ,2002(3) :16.
- [2]国家中医药管理局《中华本草》编委会中华本草精选本下册[M]. 上海 :上海科技出版社 ,1998 2062-2063.
- [3]Zhang X L ,Liu Y C. Studies and applications of Rhizoma Paridis[J]. Chin J Tradit Med Sci Technol (中国中医药科技) ,2000 ,7(5) 346-347.
- [4]陈清 ,阎姝. 重楼的药理作用及其毒性反应的研究进展[J]. 医药导报 ,2012 ,31(7).
- [5]宁远. 蒙药那如-3 中的微量元素含量测定[J]. 微量元素与健康研究 ,1999 ,16(1) :31.
- [6]唐森本. 环境化学与人体健康[M]. 北京 :中国环境科学出版社 ,1989 :124.
- [7]范子龙. 含微量元素的中药对老年病的治疗[J]. 广东微量元素科学 ,1994 ,1(6) 5 - 8.
- [8]李焘 ,屈新运 ,张序贵 ,等. 微波消解-火焰原子吸收光谱法测定 2 种不同来源重楼药材中的金属元素含量[J]. 光谱实验室 ,2011 ,28(1) :113-117.
- [9]王元忠 ,李淑斌 ,郭华春 ,等. 大百合中微量元素测定的研究[J]. 光谱学与光谱分析 ,2007 ,27(9) :1854.
- [10]吴珊 ,马宗祥 ,侯凯 ,等. 不同来源地重楼药材中 9 种无机元素含量测定[J]. 安徽农业科学 ,2010 ,38(8) :4080-4083.
- [11]虞锐鹏 ,成则丰 ,贡小清. 微波消解-原子吸收法分析并比较四种贝母中的微量元[J]. 光谱学与光谱分析 ,2007 ,27(12) :1090-1093.
- [12]韦建荣. 高效液相色谱法测定重楼提取物中偏诺皂苷类成分的研究[J]. 药物分析杂志 ,1997 ,17(3) :153-155.
- [13]陈昌祥 ,周俊 ,张玉童 ,等. 滇重楼地上部分的甾体皂苷[J]. 云南植物研究 ,1990 ,12(3) :323-329.
- [14]蒋雪 ,张经硕 ,邱敦 ,等. 重楼皂甙提取工艺优化与近红外光谱在线质量控制[J]. 分析化学 ,2006 ,34(S1) :171-174.
- [15]黄芸 ,崔力剑 ,刘伟娜 ,等. HPLC-ELSD 法分析重楼商品药材中甾体皂苷的含量[J]. 中国中药杂志 ,2006 ,31(15) :1230-1233.
- [16]刘海 ,张婷 ,陈筱清 ,等. 云南重楼的甾体皂苷类成分[J]. 中国天然药物 ,2006 ,4(4) :264-267.
- [17]中国标准出版社总编室. 中华人民共和国国家标准[M]. 北京 :国家技术监督局 ,1994 :427-430.
- [18]尹鸿 ,薛丹 ,吴梅 ,等. 川滇地区重楼商品药材质量评价[J]. 中药材 ,2007 ,30(7) :771-774.
- [19]国家药典委员会. 中华人民共和国药典(一部)[M]. 北京 :化学工业出版社 ,2010 :243.
- [20]张树平 ,陈兴福 ,杨文钰 ,等. 不同产地川泽泻中人体必需微量元素的含量分析[J]. 药物分析杂志 ,2010 ,30(7) :1213-1217.

[责任编辑:李书华]