

文章编号:1007-757X(2005)11-015-06

英语语音合成中基于约束树的音节切分算法

王永生, 柴佩琪, 曾令平

摘 要:在基于单元选择的英语语音合成中,音节切分是非常重要的一环。这主要是因为英语词汇的无限性,且存在大量的多音节词,因而许多情况下,我们不得不从音节的角度来合成语音。为此,本文提出了一种基于约束树的音节切分算法,通过应用形态规则约束、韵律规则约束及音位规则约束,从而最终切分出所有音节。在实施音位规则约束时,我们提出了一种基于音位结构树的子算法。训练样本和测试样本的切分正确率分别为99.13%和98.55%。

关键词:语音合成;音素学;音节切分;约束树

中图分类号:TP391

文献标识码:A

1 引言

语音合成常用的方法有三种^[1]:规则合成法、参数合成法和波形拼接合成法。近年来,一种基于单元选择的波形拼接法正得到越来越多的应用。在这种方法中,合成的语音单元(可以是半音素、音素、双音素、三音素、音节、短语等多种形式)是从一个预先录制的大规模自然语音库中挑选出来的。由于选取的单元都是基于自然的原始发音,它隐含了重音等韵律特征,从而保证了最终合成的语音清晰自然。

在基于单元选择的英语语音合成中,音节切分是非常重要的一环。我们知道,汉字基本上都是单音节字,并且音节总数只有一千多个^[2]。因而,只要能事先创建一个包含所有音节并且涵盖绝大多数上下文环境中的发音情况的语音库,则对于要进行合成的文本中的汉字,总能在语音库中找到相同的音节加以匹配^[3]。而英语却非如此,一方面,英语有着无限多的词汇,一部常用的词典可能动辄就会有上百万个词条;另一方面,除了单音节词外,英语中还有大量的多音节词,其音节的总数也难以统计。因而,基于存储容量和合成效率的考虑,我们不可能像汉语语音合成那样,将绝大多数英语词汇的语音预先录制下来,只能做到在语音库中包含最常用的词,并涵盖最常用的音节。这样,在合成时,对于输入文本中一个词,如果在语音库中没有选择到相同的词,则需对该词进行音节切分,然后再从音节的角度来选择;退而求其次,即使连合适的音节也没有取到,则需从音素的角度来选择,但也必须考虑它的一些上下文和韵律关系,如是否处于音节首或尾,是否处于重读音节等,以便找到最适合的音素,显然也要求事先对音节进行切分^[4]。因而音节切分是英语语音合成中非常重要的一步,其切分的正确与否,可能会直接影响到最终语音合成的质量。

基于以上原因,本文提出了一种基于约束树的音节切分算法,该方法通过对英文单词实施多层约束,从而最终正确地

切分出音节。我们用于进行音节切分的词典库中的样本有9480个,基本上涵盖了英语中最常用的一些词。为了算法规则的学习和测试,我们事先在参考Merriam-Webster词典的基础上,通过手工对其中所有词的音节进行了切分,并经过认真校对和检验,以保证所有切分百分之一百的正确,然后从9480个词中,取出7480个词用于规则学习,而剩下的2000个词用作规则测试。

本文下面的部分是这么组织的:第二部分介绍音节的结构以及组成音节的各个部分的一些约束条件;第三部分介绍利用首音最大原则切分音节核心之间的辅音丛;第四部分详细介绍基于约束树的音节切分算法;第五部分总结并指出仍需完善的地方。

2 音节结构

关于什么是音节,语言学家一直难以形成统一的意见,但从生成音系学(Generative phonology)的角度上讲,一般认为音节是根据响度(sonority)来组织音段成分的音系单位^[5]。

要切分音节,首先必须了解音节的结构。比如说,对于英文单词book,它是一个单音节词,对应的音节为/buk/,在元音/u/的前面和后面各有一个辅音,这是一个典型的音节。一般认为音节由首音(onset)和韵(rhyme)两部分组成。而韵又由核心(nucleus)和尾音(coda)组成^[5]。因而音节/buk/的结构如图1所示(其中S、O、R、N、Co分别代表音节、首音、韵、核心、尾音)。

绝大多数音节的核心是元音(vowel),但也有一些音节的核心由响音(sonorant)充当,如单词button的第二个音节/tn/,其核心就由响音/n/来充当。

核心元音或响音是音节中必不可少的成分,而首音和尾音是可选的。因而在英语中音节的基本形式主要有四种,即V、CV、VC和CVC。一些复杂的音节也是由这四类基本形式

作者简介:王永生,同济大学计算机科学与工程系,博士研究生,上海 200092

柴佩琪,同济大学计算机科学与工程系,教授,博士生导师,上海 200092

曾令平,同济大学计算机科学与工程系,上海 200092

扩展来的,如单词 abstract 中的第二个音节/str? kt/,就是由基本形式CVC扩展而来,如图2所示。为了统一起见,在本文中认为音节/str? kt/属于CVC形式,只不过其中的C不是指单个的辅音,而是指辅音丛(consonant cluster)。



图1 音节/buk/的结构

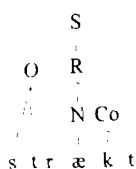


图2 音节/str? kt/的结构

了解了音节的结构以后,下面就来讨论究竟哪些音素可以充当音节的核心、首音和尾音。关于音节的组构存在多种理论,其中比较实用的一种就是响度原则。所谓响度原则(Sonority Sequencing Principle)是指一个音节从首音到核心的各个音段响度逐渐增加,而从核心到尾音的各个音段响度逐渐减弱,核心的响度最高^[7]。也就是说,语音的响度是分等级的,其等级划分如表1所示^[7]。

表1 语音的响度等级

大类	小类	所包含的音素	响度等级
响音 (sonorant)	元音(vowel)	所有元音	6
	滑音(glide)	w, j	5
	流音(liquid)	r, l	4
	鼻音(nasal)	m, n, ŋ	3
阻塞音 (obstruent)	摩擦音/塞擦音 (fricative/affricate)	v, e, z, ((浊音) f, (, s, (, h(清音) t, d((塞擦音)	2
	爆破音(plosive)	b, d, g(浊音) p, t, k(清音)	1

下面来看看如何根据响度原则确定一个音节的核心、首音及尾音。

2.1 音节核心

由于在一个音节中核心是必不可少的,那么就先来研究一下音节的核心。在本文中,我们把/ə/以及一些双元音与单元音/ə/的组合,如/eia/、/aia/、/auə/、/oia/及/əua/看成是英语的基本音,这样一来,加上28个辅音、26个基本元音(包括20个单元音和6个双元音),在本文中,我们认为英语中的基本音有54个^[8],如单词coward[kauəd],我们就认为它是一个单音节词。

既然核心是音节必不可少的组成部分,那么如果能先正确地找出一个单词音标中的所有音节核心,那么有多少个核

心,就必然有多少个音节。如上所述,音节核心可能有两种:

(1) 元音

无论是单元音、长元音、双元音,还是三元音,我们都认为它是一个基本元音,如but[bʌt]、seet[si:t]、show[ʃəu]及shower[ʃauə]中的/ʌ/、/i:/、/əu/及/auə/均是各自音节的核心。

(2) 音节辅音

绝大多数音节的核心都是由元音来充当,但也有少数音节的核心由响辅音/m/、/n/和/l/来充当。如burden[ˈbɜːdn]的第二个音节/dn/,其音节核心为响鼻音/n/、/d/为该音节的首音。

至所以将/dn/看成一个独立的音节,而不是前一个音节的尾音,从响度原则的角度来看,也是不难理解的,在长元音/ə:/处响度达到最高,按照表1的定义,其响度为6,而它后面的爆破音/d/的响度为1,但/d/后面的鼻音/n/的响度为3,如果将整个看成一个音节,显示违背了响度原则。

在英语中,由响音充当音节核心的情况主要有4种^[9],如表2所示。

表2 响音充当音节核心

阻塞音+流音l	如:title[ˈtai-tl]、ankle[ˈ(ɪ)k-l]
阻塞音+鼻音m	如:optimism[ˈ(p-ti-mi-zm)]
阻塞音+鼻音n	如:burden[ˈb(ɜː-dn)、basin[ˈbei-sn]
鼻音m+流音l或 鼻音n+流音l	如:mammal[ˈm(-ml)、 national[ˈn(-(n)l]

在音系学中,像流音l、鼻音m和n的这些能充当音节核心的响音统称为音节辅音(syllabic consonants)。

流音/l/和鼻音/m/、/n/可以作为音节辅音来充当音节核心,但问题是这三个音素什么时候是一般的辅音(是首音或尾音的组成部分),什么时候又是音节核心呢?这三个音素要成为音节辅音必须满足下列二个条件:

✓在其前面有一个阻塞音(这个阻塞音也就成为该音节的首音)或是在流音/l/前是鼻音/m/或/n/

✓在l、m、n后紧跟着的一个音素不是元音

因而虽然costly[ˈk(ɒ)stli]中也存在由阻塞音/t/+流音/l/的组合,但由于在/l/后有一个元音/i/,因而在这里流音/l/不是音节辅音,而只是音节[li]的首音。

而decent[ˈdi:snt]中的鼻音/n/完全符合上述两个条件,因而它是音节辅音,也就是说/n/是第二个音节的核心,该音节的首音为/s/,尾音为/t/。

2.2 音节首音

英语中音节的首音可以有0、1、2或3个辅音,最多不超过3个辅音,而尾音可以由0到4个辅音组成^[10]。针对首音中辅音个数的不同,下面来分别加以讨论。

(1) 1个辅音

在所有单辅音中,除了鼻音ŋ外,其余我们都可以轻而易举地找出以它们为首音的音节例子。而鼻音ŋ是否可以直接用作音节的首音呢?显然不能,因为在英语中,鼻音ŋ只可以紧跟在双唇塞音k和g的后面^[9],如inker[ˈɪŋkɪ]和coming

['kʌmiŋ]。

(2) 2个辅音

首音由两个辅音组成的情况比较复杂,究竟哪些辅音组合可以用作音节的首音呢?可以根据最小响度距离原则(Minimum Sonority Distance Principle)^[11]来解决这个问题。所谓最小响度距离原则,不妨以首音由两个辅音组成的情况为例,就是说,只有当第二个辅音的响度与第一个辅音的响度之差大于等于2时,由这两个辅音组成的辅音丛才有可能作为音节的首音。这样一来,符合这一原则的双辅音组合就只有以下4种情况^[9],如表3所示。

表3 可以充当音节首音的双辅音组合

塞擦音+流音	pl pr tr kl kr fl fr θr sl fr bl br dr gl gr
塞擦音+滑音	pj tj tw kj kw [j θj (w sj sw hj bj dj dw gj gw vj)
鼻音+滑音	mj
鼻音+流音	nj

除了上述4种情况外,还有一种特殊情况,就是以清摩擦音/s/为首的双辅音丛,如stay[stei]。/st/作为音节的首音显示违背了最小响度距离原则,因为/s/的响度为2,而/t/的响度为1,后者与前者之差为-1,但它确实是可以作为音节的首音。这类情况虽然特殊,但仍有规律可循,只可能存在两种情况,即:

s+清爆破音和s+清摩擦音

其可能的组合也只有sp、st、sk和sf这4种情况^[9]。

(3) 3个辅音

三辅音丛作音节首音的情况在英语中占的比重不大,且均以s为首。这样一来,可能的组合就会很少,如上所述,s后面只能跟清音,而第二个辅音与第三个辅音的组合又必须符合最小响度距离原则,因而符合条件的三辅音丛必须满足^[12]:

s+(p, t, k) + (w, j, r, l)

所以可能的组合只有:spl spr spj str stj stw skl skr skj skw

如:splay [splei]、stray [strei]、scrape [skreip]、spew [spju:]、stew [stju:]、skew [skju:]、squish [skwi:]。

3. 音节尾音

一个音节可以没有尾音,也可以有一个由1个、2个、3个或4个辅音组成的尾音。与首音相比,对尾音的约束条件较为宽松,因而它的组合情况就比较复杂。显然尾音的构成不符合最小响度距离原则,但基本上符合响度原则。

3 利用首音最大原则切分音节核心之间的辅音丛

上面我们介绍了音节的结构及组成音节的三个部分,即核心、首音和尾音。了解了这些,还不足以让我们能很容易地切分出音节。比如说,access[ækses]是一个双音节词,这两个

音节的核心是很容易确定的,分别是/æ/和/e/,第一个音节的首音和第二个音节的尾音也是显而易见的,问题在于两个音节核心之间有两个辅音/k/和/s/,而这两个辅音存在三种可能的切分,分别如图3的(a)、(b)和(c)所示(其中w代表词)。

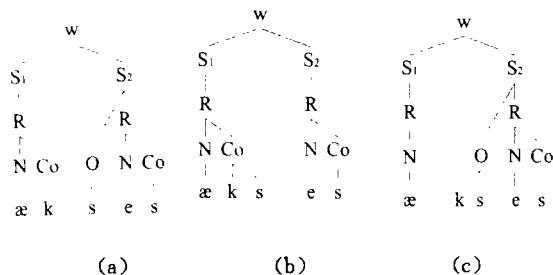


图3 [(kses)]的三种可能的音节切分

很显然,第三种情况将/ks/作为第二个音节的首音是不对的,因为它违背了最小响度距离原则。而(a)和(b)中究竟哪一个才是正确的切分呢?像这类音节核心之间的辅音(inter-vocalic consonant)的划分,正是音节切分最大的难点。可以利用首音最大原则^[13](Maximum Onset Principle)来解决这一问题,所谓首音最大原则,是指对于音节核心之间的辅音,总是分配尽可能多的辅音作为后一个音节的首音。下面就根据两个音节核心之间辅音个数的不同分别讨论如何利用首音最大原则对他们进行切分。

3.1 一个辅音

如果两个音节核心之间只有一个辅音,则根据首音最大原则,该辅音应该作为后一个音节的首音。只有一个辅音例外,即鼻音ŋ,如果在两个音节核心之间只有一个鼻音ŋ,则它应该作为前一个音节的尾音,如singer['sɪŋə]。

3.2 两个辅音

这类情况比较复杂,在所有用于测试的7480个词中,存在这一情况的有2846个词,比例高达38%,而在这2840个词中,存在3287个核心间双辅音。

这类情况可用模式 $N_1C_1C_2N_2$ 来表示,如果用 S_1 和 S_2 分别表示相邻的两个音节,则 N_1 、 N_2 分别表示 S_1 和 S_2 的核心, C_1C_2 表示核心 N_1 和 N_2 之间的两个辅音。根据首音最大原则,切分 C_1C_2 的算法如下:

```

If  $C_1C_2$  是符合首音条件的双辅音丛 Then
 $C_1C_2$  为  $S_2$  的首音
Else
If  $C_2$  是符合首音条件的辅音 Then
 $C_2$  为  $S_2$  的首音,  $C_1$  为  $S_1$  的尾音
Else
 $C_1C_2$  为  $S_1$  的尾音
End If
End If

```

就以absense['æbsəns]为例,很显然,该词为双音节词,两个音节核心分别为/æ/和/ə/,现在要考虑双辅音/bs/的划分,首先判断/bs/是否可作为第二个音节的首音,根据最小响度距离原则,/bs/不是一个可作首音的辅音组合,则再判断/s/

是否可用作后一个音节的首音,显然是可以的,因而最终的切分结果是/b/为第一个音节的尾音,而/s/为第二个音节的首音。

3.3 个及3个以上辅音

这类情况虽然不多但较复杂,在所有用于测试的7480个词中,存在这一情况的有924个词,比例达12.4%,而在这924个词中,存在942个核心间三个或三个以上的辅音丛。

这类情况可用模式 $N_1C_1C_2 \dots C_nN_2$ 表示,根据首音最大原则,其切分算法如下:

If $C_{n-2}C_{n-1}C_n$ 是符合首音条件的三辅音丛 Then

$C_{n-2}C_{n-1}C_n$ 为 S_2 的首音, $C_1 \dots C_{n-3}$ 为 S_1 的尾音

Else

If $C_{n-1}C_n$ 是符合首音条件的双辅音丛 Then

$C_{n-1}C_n$ 为 S_2 的首音, $C_1 \dots C_{n-2}$ 为 S_1 的尾音

Else

If C_n 是符合首音条件的辅音 Then

C_n 为 S_2 的首音, $C_1 \dots C_{n-1}$ 为 S_1 的尾音

Else

$C_1 \dots C_n$ 为 S_1 的尾音

End If

End If

End If

虽然首音最大原则是一种通用的原则,在绝大多数情况下是有效的,但也存在例外,如 *accustom* [ə'kʌstəm], 如果按照首音最大原则,其音节切分应为 [ə-'kʌstəm], 因为 /st/ 是一个符合条件的音节首音,然而正确的切分却是 [ə-'kʌstəm]。经过统计分析,这类例外较多地出现在以 /s/ 为首、且不符合最小响度距离原则的情况下,即针对组合 /sp/、/st/、/sk/、/sf/、/spl/ 等,如果这类辅音组合出现首音节中,或出现在前缀后的第一个音节中,则可应用首音最大原则,将他们视为音节的首音,而在其他情况下,应该被切分开,即 /s/ 作为前一个音节的尾音,而剩下的部分作为后一音节的首音。

4 基于约束树的音节切分算法

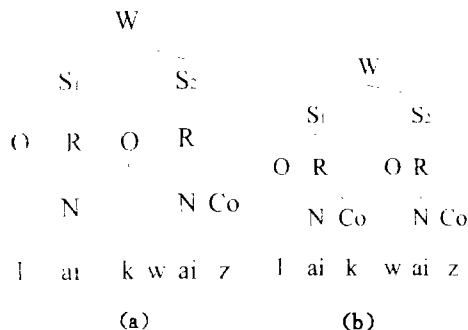


图4 [laikwaiz] 的两种音节切分

上一部分我们主要从音系学的角度讨论了音节核心和音节核心之间的辅音的切分,但仅仅凭借这些知识来切分音节,许多情况下还是难以奏效的。如 *likewise* ['laikwaiz], 这个词的切分关键就在于元音 /ei/ 和 /ʌ/ 之间的 /kw/ 的归属,根据首音

最大原则, /kw/ 应作为第二个音节的首音,即如图4的(a)所示。但实际上这个词的第二部分,即 *wise* 为其后缀,组成这个词的两部分的接合处就是音节的边界,因而正确的切分应如图4的(b)所示。

由此,我们想到,在使用音系规则前,最好能先应用其他一些规则,如形态规则,以尽量降低音节切分的错误率。为此,我们提出了一种基于约束树(Constraint Tree)的音节切分算法,通过应用多层次的规则约束,从而最大限度地提高音节切分的正确率。

为了清楚地说明这一算法,不妨以单词 *congratulation* [kən,grætju'leɪʃən] 为例,该词的约束树如图5所示。

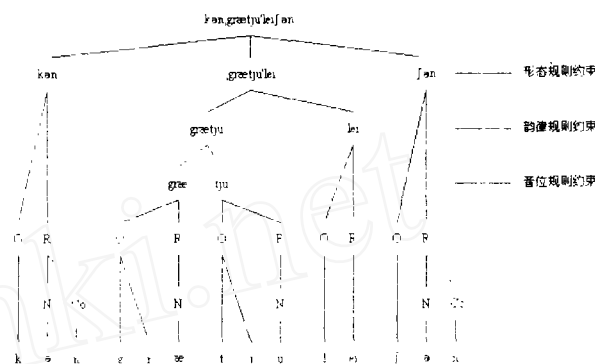


图5 [k(n,gr? tju'lei(n)] 的约束树

由图5可以看出,在整个约束树中,共实施了三层规则约束,下面分别对这三层规则约束加以说明。

4.1 形态规则约束(Morphological Rules Constraint)

所谓形态规则约束,是指从词的内部结构和支配构词的一些规则来约束音节的切分。当然本文并不会深入到词素的程度来研究音节切分,只是应用形态学的一些知识来辅助音节切分。比如说,研究发现,对于一些复合词,构成词的两部分的接合处往往就是音节的边界;而对于有前缀和后缀的词,其前缀或后缀与词干的接合处也往往就是音节的边界。因而,如果能事先对复合词和有前缀或后缀的词进行预切分,则可大大简化音节的切分,并可提高切分的正确率。下面分别针对这两种情况进行讨论。

(1) 复合词

复合词又可分成两类:

(a) 有连字符的复合词

对于有连字符的复合词,组成复合词的两部分的切分往往不受音系规则(如响度原则、首音最大原则等)的约束,复合词两部分的接合处往往就是音节的边界。如 *make-up* ['meik-ʌp], *self-aware* [self-ə-'weə] 等。(b) 无连字符的复合词

对于无连字符的复合词,组成复合词的两部分的切分往往要受音系规则的约束,如 *make-up* 有时也写成 *makeup*,但此时其切分就为 ['mei-kʌp], 而非 ['meik-ʌp]; 再比如 *pineapple* ['pai-næ-pl], 而非 ['pain-æ-pl]。

(2) 有前缀或有后缀的词

对于大多数有前缀或后缀的词,其前缀或后缀与词干的接合部往往就是音节的边界,但它们也要受音系规则的约束。

如expert['ek-spə:t], inaccessible[i-naek-'se-sə-bl]及roughly['r(-fli), helpless['hel-plis]等。因而在进行切分时必须充分考虑前缀或后缀与词干接合处的辅音的切分,如图5所示,congratulation 就有一个前缀con 和一个后缀ion,因而通过正确地识别其前后缀,就可以将它的音节预切分成三部分,即/kən/, /graetju'lei/和/f(n/。

4.2 韵律规则约束(Metrical Rules Constraint)

我们知道,英文单词都是由一些韵律单元(Rhythmic Units)所组成,所谓韵律单元,就是指组成词的音节中的重读音节、次重读音节和非重读音节。由一个重读音节加上一个或多个非重读音节构成的单位一般称为韵脚(foot),它用于表示一个词的韵律结构^[14]。

英语中韵脚是左支配的(left-dominant),也就是说,在一个韵脚中,第一个音节总是重读音节或次重读音节(也称强音节),其后面的音节为非重读音节(也称弱音节)。如apparatus[,aepə'reitəs]中就有两个韵脚,分别为[aepə]和[rei-təs],这两个韵脚都是由两个音节组成,且最左边的音节都是强音节,而后一个音节是弱音节。而在这两个韵脚中,前一个韵脚的首音节是重读音节,后一个韵脚的首音节是次重读音节,因而前一个韵脚也称为强韵脚,相应的,后一个韵脚则称为弱韵脚。

显然,韵律特征对于音节切分来说,是一种非常重要的区别性特征(distinctive feature),利用这个特征,可以再次对一个词的音节进行预切分。比如说,根据韵律特征,就可以将词apparatus 预切分成两部分,即[aepə]和[reitəs]。同时,通过应用韵律规则,也可以取得音节的一些韵律特征,这些韵律特征将在以后的单元选择和波形拼接中用到。

4.3 音位规则约束(Phonotactic Rules Constraint)

关于音位规则约束,在第二部分和第三部分已详细介绍过。为了更好地应用音位规则约束,我们这里设计了一个音位结构树,作为整个约束树的一个子树,以用于对经过形态规则约束和韵律规则约束后切分出来的音节组合按照音位规则进行最后的切分。就以obstacle['(bst(kl)]为例,来说明该部分的切分算法。

(1) 首先生成一个叶子结点均为空的音位结构树;然后找出所有音节的核心,包括元音核心和音节辅音核心,并将找出的音节核心填充到各自音节的核心结点中;再将第一个音节核心前的辅音填充到该音节的首音结点中;将最后一个音节核心后的辅音填充到该音节的尾音结点中;如果某个音节无首音或尾音,则在对应的结点中填充"/"。这一步称为音位结构树的初始化,经过初始化后的音位结构树如图6所示。

(2) 从右往左,首先切分 s_3 与 s_2 之间的辅音,根据上一部分叙述的算法,/k/应该是 s_3 的首音,将它填充到 s_3 的首音结点中。

(3) 再考虑 s_2 与 s_1 之间的辅音的切分,即/bst/,根据上一部分叙述的算法,/st/应该是 s_2 的首音,/b/应该是 s_1 的尾音,则将/st/和/b/分别填充到 s_2 的首音结点和 s_1 的尾音结点中。

由于已遍历到了 s_1 ,所以,所有音节切分完毕。最终生成的音位结构树如图7所示。

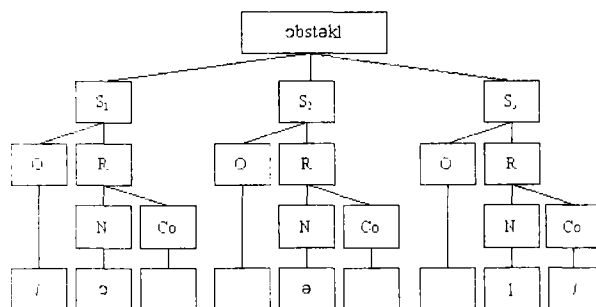


图6 经过初始化的音位结构树

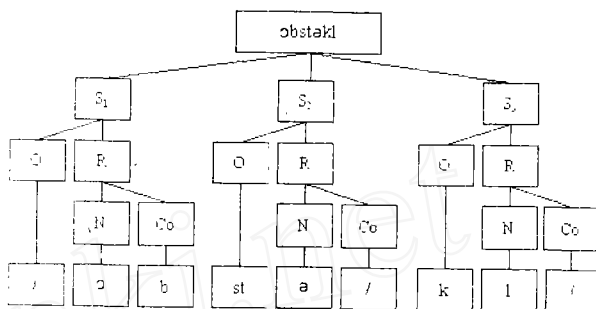


图7 全部音节切分完毕后的音位结构树

具体的算法可以描述如下:

找出所有音节的核心,包括元音核心和音节辅音核心,并填充音节核心结点、第一个音节的首音结点及最后一个音节的尾音结点

For I=音节的个数 To 2 Step-1

切分音节 s_i 的核心与音节 s_{i-1} 核心之间的辅音,并将切分出来的 s_i 的首音填充到其首音结点,将 s_{i-1} 的尾音填充到其尾音结点

Next I

5 结论

由于英语有着无限多的词汇,其中又有大量的多音节词,我们不可能、也没有办法预先录制涵盖所有词汇的语音,所以许多时候我们不得不从音节的角度来进行拼接,因而音节切分是基于单元选择的波形拼接语音合成中非常重要的一环。本文提出了一种基于约束树的音节切分算法,通过应用形态规则约束、韵律规则约束及音位规则约束,从而最终切分出所有音节。在实施音位规则约束时,我们提出了一种基于音位结构树的子算法。我们用于音节切分的词典库中的样本数目为9480个,在这9480个词中,我们取出7480个词用于规则学习,而剩下的2000个词用作测试。最终的结果为,训练样本的切分成功率为99.13%,而测试样本的切分成功率为98.55%。

很显然,还是存在极少数的切分错误的情况,训练样本集中有65个词切分错误,而测试样本集中有29个词切分错误,经过对这些词逐一加以分析,发现许多词的音节之所以切分错误,是因为他们违背了我们算法所依据的一些原则,如menu['menju],根据我们算法,其切分应为['me-nju],而实

际上正确的切分应为['men-ju],这显然违背了首音最大原则,因为/nj/在绝大多数情况下都是作音节首音的。既然仍存在一些切分错误的情况,如何进一步消除这些错误,是我们仍需继续努力的地方。

参考文献:

- [1] Chen Xiaolin. The development of Chinese text-to-speech system [J], Computer & Communication, 1998, 3: p1~4
- [2] 张雄伟, 陈亮. 现代语音处理技术及应用[M]. 北京: 机械工业出版社, 2003
- [3] 吴晓如. 多语种语音合成中的关键技术研究[D]. 安徽: 中国科技大学, 2002
- [4] Jon Rong-Wei Yi. Corpus-Based Unit Selection for Natural-Sounding Speech Synthesis [D]. USA: MIT, 2003
- [5] Blevins, J. The syllabic in phonological theory [A]. J. Goldsmith. The Handbook of Phonological Theory [C]. Cambridge: Blackwell Publishers, 1995, P206-244
- [6] Fudge, E. Syllables [J]. Journal of Linguistics, 1969, 5: P253-287
- [7] Selkirk, E. On the major class features and syllable the-

- ry [A]. M. Aronoff and R. T. Oehrle. Language Sound Structure [C], Cambridge: MIT Press, 1984, F105-136
- [8] 李桂荣. 英语语流中的语音语调 [M]. 北京: 经济管理出版社, 2003
- [9] Dan Mateescu. English Phonetics and Phonological Theory, <http://www.unibuc.ro/eBooks/filologie/mateescu/cuprins.htm>, 2003
- [10] Mackay, I. R. A. Phonetics: The Science of Speech Production [M], Boston: Little and Brown., 1987
- [11] Harris, James W. Syllable structure and stress in Spanish: A nonlinear analysis [M]. Cambridge: MIT Press, 1983
- [12] Robert Mannell: an Introduction to Phonetics and Phonology, <http://www.ling.mq.edu.au/units/ling210-901/main/home/index.html>, 2004
- [13] Selkirk, Elisabeth O. The syllable [A]. Harry van der Hulst, Norval Smith. The Structure of Phonological Representations [C]. Dordrecht: Foris. 1982, P337-383
- [14] Levin, Juliette. A metrical theory of syllabicity [D]. USA: MIT, 1985

(收稿日期: 2005-5-17)

(上接第47页)

目前, 无线视频传输大多利用 2.4GHz 的免申请频段, 这个频段属于微波, 利用微波传输视频信号, 信号好, 干扰小, 图像无失真, 可以传输彩色图像。这个频段也可以无线传输声音, 可以传输立体声。

本系统采用了上海金鑫电子有限公司的 2.4G 无线影音传输模块。利用 2 对该模块可以实现室外机向室内机方向的视频信号的无线传输和音频信号传输, 以及室内机向室外机音频信号的传输。该模块采用了先进的微波无线电传输技术, 可以高保真的无线传输视频信号和音频信号。

该模块主要有以下几方面特点:

(1) 良好的微波抗干扰性能。该模块音/视频信号采用频率极高的微波传输方式, 其载波频率高达 2.4G。因此可以有效地避免 900MHz 等频段拥挤不堪的无线电信号或各类电火花及家用电器的干扰, 同时也极大程度地避免了本机发射频率对其他接收设备的干扰。

(2) 优良的音/视频接收品质。该模块为提高音/视频接收品质, 采用了性能优越的调频方式。音/视频具有宽频特性, 且信噪比高。使用监控摄像头传输信号, 也有极佳的分辨力。(3) 该模块采用单一电源 +5V 供电, 便于使用。

(4) 广泛的应用扩展空间。

(5) 如果发射信号源是无声的, 则音频线可免接。

为了避免产生门铃误响的现象, 我们选择了该模块, 可以注意模块上设置了 4 个频道可供选择, 每个频道相隔 18MHz,

避免了互扰现象。这样同时也保证了声音和图像的清晰接收。

5 结束语

本文所设计的无线可视双向对讲门铃, 可以有效的解决有线门铃带来的布线问题。利用了无线数字传输技术, 使得无线控制得以实现, 利用无线音视频传输技术, 实现了室内对门外情况的监视, 以及两者的对讲。另外, 本系统采用了模块化设计思想, 便于进行项目扩展。

参考文献:

- [1] 丁立波, 江小华, 张河. 基于无线数字传输技术的数据采集系统设计 [J]. 半导体技术, 2003(02)
- [2] 夏方林. 基于 AT89C2051 的单户可视对讲门铃室内分机设计 [J]. 微计算机信息, 2004(10)
- [3] 沈兰荪, 田标. 无线视频传输技术的发展 [J]. 电子技术应用, 2001(01)
- [4] SRWF-1 型微功率无线数传模块使用说明书 [Z], 上海桑锐电子科技有限公司, 2004
- [5] 2.4GHz 无线影音传输模块说明书 [Z], 上海金鑫电子有限公司, 2004

(收稿日期: 2005-5-11)

ISSN 1007-757X

Microcomputer Applications

Monthly
(Since 1985)Wu Qidi
Editor-in-Chief
Vol. 21, No. 11 (General No. 151)
November 2005

Research & Design

Modeling and Analysis of Network Traffic Prediction Using Autoregression and Support Vector Machine Based on Byte Length in Data Packets P(1)*Wang Jinghua* (Department of Electrical Engineering, Shanghai Jiaotong University, Shanghai 200240);
He Di (Department of Electronics Engineering, Shanghai Jiaotong University, Shanghai 200240)**Abstract** This paper presents the results of two methods used in predicting network traffic for intrusion detection. These methods are linear prediction using autoregression, and nonlinear prediction using support vector regression. This paper will give details on the effectiveness of detecting attacks relative to each prediction method. It was found that the SVM for regression improved the detectability of the attacks and lowered the false alarm rates compared with AR model. In addition, the AR model is less computationally expensive when compared with SVM regression.**Keywords** intrusion detection techniques NIDS Auto Regression SVM SVC**Formal Method Integrating Petri Nets and B** P(4)*Jiang Mengzhi Zhang Guixu* (Department of Computer Science, East China Normal University, Shanghai 200062)**Abstract** In this paper, a formal method with the capability of specifying distributed system and real-time system is presented. named Petri-B Nets, which integrates Petri Nets with B Method. It not only reserves the characteristics of Petri Nets, but also contains the advantages of B Method. In this paper we define the Petri-B Nets and use it to describe a Seat-Reservation System.**Keywords** formal method Petri Nets B Method Petri-B Nets**Design of VMI's Model in Fishery Industry** P(7)*Cui Fang Chen Ming* (College of Information, Shanghai Fisheries University, Shanghai 200090)**Abstract** Vendor Managed Inventory(VMI) is where the manufacturer assumes responsibility for the atomic replenishment of a customer's inventory. A well-managed VMI program delivers benefits to both the buyer and the supplier. Considering the feature of fishery goods and inventory management, we analyze the VMI's basic feature and acquire the model of the fishery industry.**Keywords** VMI supply chain the fishery industry inventory**A Semantic-Coding based Watermarking Algorithm** P(10)*Li Jin* (Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200030)**Abstract** This article presents a semantic-coding-based blind watermark detection algorithm. At first, the algorithm exploits the two-dimensional barcode technique to encode watermark information, then utilizes chaotic mapping to make watermark in confusion, and at last embeds the watermark into the DCT coefficients of the carrier image. This algorithm is robust not only in the general lower level of DSP, but also in the upper level of semantic, by taking the advantage of the self-checking ability of two-dimensional barcode. With its high capacity for information, semantic and self-checking characteristics, this practical algorithm can implement real-time detection on the network.**Keywords** watermark data hide semantic-coding two-dimensional barcode chaos**Research of Binary Trade-mark Image Retrieval Method Based on NMI Invariable Feature** P(13)*Wang Zhenhai* (Department of Information, Linyi Normal University, Linyi 276005)**Abstract** Trade-marks' retrieval has obtained more and more attention in recent research on content-based management and utilization of image database system. This paper proposes an image retrieval method based on NMI(Normalized Moment of Inertia) invariable feature for binary trade-mark images, which is simple with high efficiency, and has the ability of anti-geometrical distortion while robust to some image-processing such as filtering and smoothing. Experimental results show that it can be used for trade-mark image retrieval.**Keywords** trade-mark NMI feature image retrieval**A Syllabification Algorithm Based on Constraint Tree in English Speech Synthesis** P(15)*Wang Yongsheng Chai Peiqi Zeng Lingping* (Department of Computer Science and Engineering, Tongji University, Shanghai 200092)**Abstract** Syllabification is a fundamental step in English Speech Synthesis based on unit selection. The main reason is that there are infinite words in English, many of which are polysyllabic. So in many cases we have to do the synthesis based on syllables. This paper proposes a syllabification algorithm based on constraint tree, in which morphological rules constraint, metrical rules constraint and phonotactic rules constraint are imposed. About phonotactic rules constraint, a sub-algorithm based on phonotactic tree is presented. The overall accuracies of training and testing sets are 99.13% and 98.55% respectively.**Keywords** Speech Synthesis Phonology Syllabification Constraint Tree

Development & Application

GSM-based Wireless Data Exchange in Aquatic Breeding P(21)*Shi Bing* (Computer Department, Jiangsu Polytechnic University, Changzhou 213016)**Abstract** This paper introduces a wireless data-transmission system based on GSM for Aquatic Breeding monitoring. The