

一种实时基音检测算法

王都生, 铁满霞, 樊昌信

(西安电子科技大学信息科学研究所, 西安 710071)

摘 要: 为了克服目前诸多基音周期检测算法迟延较大的缺点, 本文提出了一种基音周期的实时检测算法. 实验结果表明: 该算法可在当前帧内实时提取精确的基音周期估计值, 减少了传统的自相关基音检测算法因平滑处理而带来的编码迟延.

关键词: 基音检测; 实时性; 语音编码

中图分类号: TN912.13 **文献标识码:** A **文章编号:** 0372-2112 (2000) 10-0002-03

A Real Time Pitch Detection Algorithm for Speech Signal

WANG Du2sheng, TIE Man2xia, FAN Chang2xin

(Information Science Institute, Xidian University, Xi. an 710071, China)

Abstract: A real time pitch detection scheme is proposed in this paper in order to reduce the delay of the existent pitch detection algorithms. The experimental results show that the new scheme can obtain exact pitch level within the current frame and reduce the delay of the existent autocorrelation pitch detection algorithms using smooth processing.

Key words: pitch detection; real time; speech coding

1 引言

在用传统的自相关基音检测法估计基音周期时, 判决错误的可能性较大, 一般采用谱平整的预处理^[1]和基音平滑后处理^[2]等技术以改善基音估计的性能. 由于平滑处理需要缓存至少一帧的语音信号样点, 因而不可避免地带来了额外的编码迟延, 这在许多实际应用中是非常不利的. 为此本文提出了一种新的归一化自相关函数基音检测算法. 该算法采用整数基音周期初估、分数基音周期精估及基音周期倍数检验等处理步骤, 可在当前帧内实时提取精确的基音周期估计值, 减少了传统的自相关函数法提取基音周期过程中因平滑处理而带来的编码迟延.

2 基音实时检测算法

新的基音检测算法的原理框图如图 1 所示, 主要由以下 4 个步骤组成:

- (1) 根据 1000Hz 的低通滤波输出语音信号, 计算整数基音周期初始估计值 P_1 ;
 - (2) 对 500Hz 的带通滤波输出语音信号进行基音周期精确估计值 P_2 的求取;
 - (3) 根据长时平均基音周期 P_{avg} , 由 1000Hz 的低通滤波残差信号确定最终基音周期 P_3 ;
 - (4) 进行长时平均基音周期 P_{avg} 的更新.
- 下面详细叙述每一步算法的具体运作过程.

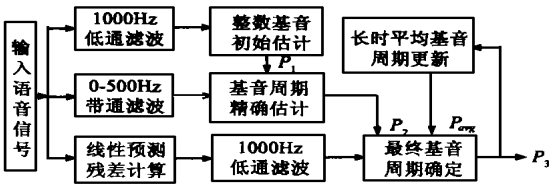


图 1 基音周期实时检测算法框图

2.1 整数基音周期初始估计

对截止频率为 1000Hz 的低通滤波输出语音信号进行归一化自相关函数 $r(S)$ 的计算, 使 $r(S)$ 为最大的 S 值即是整数基音周期 P_1 . 本文采用的归一化自相关函数 $r(S)$ 定义如下:

$$r(S) = \frac{C_S(0, S)}{\sqrt{C_S(0, 0) C_S(S, S)}}, S = 40, \dots, 160 \quad (1)$$

其中: $C_S(m, n) = \sum_{k=-\delta S/28-80}^{-\delta S/28+79} S_{k+m} S_{k+n}$, m, n 均为整数. $\delta S/28$ 表示不大于 $S/2$ 的最大整数.

在整数基音周期计算过程中, 式(2)中的时标 k 确保基音分析窗围绕它的中心点两边对齐. 初始基音周期估计的范围为 $P_1 \in (40, \dots, 160)$, 在最后的基音周期计算确定中会将基音周期范围扩展到 $(20, \dots, 160)$.

2.2 基音周期精确估计

精估基音周期是在 0-500Hz 带通滤波器的输出信号段进行的. 在精估基音周期时, 考虑两个基音周期候选者, 即当

前帧的整数基音周期 P_{l_cur} 和上一帧的整数基音周期 P_{l_pre} . 对于每个候选者, 用式(1)进行候选者 5 范围内的整数基音周期搜索, 找出最佳的整数基音周期. 然后围绕该最佳整数基音周期进行分数基音周期估计, 从而将产生两个分数基音周期候选者及其对应的归一化自相关值, 对应较高归一化自相关值的分数基音周期候选者被选为基音周期精确估计值 P_2 .

其中分数基音周期估计是利用一个插值公式来提高输入基音周期值的精度, 步骤如下:

$$\$ = \frac{c_r(0, T+1)c_r(T, T) - c_r(0, T)c_r(T, T+1)}{c_r(0, T+1)[c_r(T, T) - c_r(T, T+1)] + c_r(0, T)[c_r(T+1, T+1) - c_r(T, T+1)]} \quad (3)$$

其中 $c_r(m, n)$ 由式(2)定义.

自相关值由式(4)给出:

$$r(T+\$) = \frac{(1-\$)c_r(0, T) + \$c_r(0, T+1)}{\sqrt{c_r(0, 0)[(1-\$)^2c_r(T, T) + 2\$(1-\$)c_r(T, T+1) + \$^2c_r(T+1, T+1)]}} \quad (4)$$

式(3)和(4)分别得到分数偏值及分数基音周期对应的归一化自相关值.

2.1.3 最终基音周期确定

最终基音周期的计算是在 1000Hz 的低通滤波残差信号段进行的.

首先用式(1)完成 P_2 5 范围内的整数基音周期搜索(其中 P_2 是四舍五入的整数), 接着围绕搜索到的最佳整数基音周期进行分数基音周期估计, 将产生最终的基音周期 P_3 及其对应的归一化自相关值 $r(P_3)$.

若 $r(P_3) \geq 0.6$, 则在残差信号段进行基音周期倍数检验, 其中用 P_3 作为候选基音周期, 倍数检验门限 $D_{th} = 0.75(P_3 F 100)$, $D_{th} = 0.5$ (其它情况).

否则, 在输入的语音信号段围绕 P_2 进行分数基音周期估计, 将产生新的 P_3 和 $r(P_3)$. 若 $r(P_3) < 0.55$, 则 P_3 由长时平均基音周期 P_{avg} 代替; 否则, 在输入的语音信号上进行基音周期倍数检验, 其中用 P_3 作为候选基音周期, 倍数检验门限 $D_{th} = 0.9(P_3 F 100)$, $D_{th} = 0.7$ (其它情况).

最后, 若 $r(P_3) < 0.55$, 则 P_3 由 P_{avg} 代替.

上述最终基音周期确定过程中使用基音周期倍数检验过程, 是为了寻找实际基音周期的倍数或分数, 确定更准确的基音周期. 该倍数检验过程叙述如下:

该检验过程的输入参数为一段语音信号、一个候选基音周期 P 、一个加倍检验门限 D_{th} , 输出为检验的基音周期 P_c 及其对应的 $r(P_c)$.

首先围绕基音周期 P 做分数基音周期估计, 产生较精细的 P_c 及其对应的 $r(P_c)$.

接着在 $(P_c/k) \in [20, 200]$ 且 $k = 8, 7, \dots, 2$ 的条件下, 找满足 $r(P_c/k) > D_{th}(P_c)$ 的最大 k 值. 其中 $r(P_c/k)$ 分两步计算:

(1) 围绕 P_c/k 做分数基音周期估计, 产生 P_k .

(2) 若 $P_k < 30$, 则进行加倍验证. 如果找到这样一个 k 值, 那么围绕基音周期 P_k 做分数基音周期估计, 产生新的 P_c 及其对应的 $r(P_c)$.

最后, 若 $P_c < 30$, 则进行一次加倍验证.

加倍验证用于预防突发的短基音周期值的出现. 对于输入 P 和 $r(P)$, 加倍验证步骤返回 $r(P)$ 和 $r(2P)$ 的较小者, 其

首先将输入基音周期值(如果是分数)四舍五入到最近的整数值. 假设该整数有 T 个样点值, 插值公式会认为 $r(S)$ 在 T 和 $T+1$ 之间存在一最大值, 因此, 需要计算 $c_r(0, T-1)$ 和 $c_r(0, T+1)$, 并进行比较以确定最大值是落在区间 $(T, T+1)$, 还是落在区间 $(T-1, T)$. 如果 $c_r(0, T-1) > c_r(0, T+1)$, 则最大值可能落在区间 $(T-1, T)$, 那么基音周期 T 在插值之前应减去 1. 分数偏值 $\$$ 由插值公式计算如下:

中 $r(2P)$ 是围绕 $2P$ 进行分数基音周期估计确定的.

2.1.4 平均基音周期更新

本文采用了一个简单的平滑过程对长时平均基音周期 P_{avg} 进行更新.

如果 $r(P_3) > 0.8$ 且 $G > 40\text{dB}$ (G 为本帧增益), 则 P_3 存入一个包含最近三个较强基音周期值 $P_i, i = 1, 2, 3$ 的缓冲器; 否则, 缓冲器内的三个基音周期值按照式(5)进行更新:

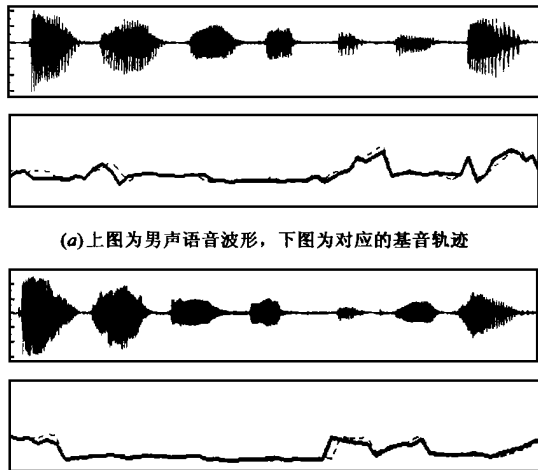
$$P_i = 0.95P_i + 0.05P_{\text{default}}, i = 1, 2, 3 \quad (5)$$

其中 $P_{\text{default}} = 50$ (样点).

平均基音周期输出为缓冲器中三个值的中间值.

3 实验及其结论

运用本文提出的基音实时检测算法对一段男女声/大家都说普通话0的语音信号进行基音周期估计, 其结果如图2所示. 图2(a)、(b)下图中的实线为采用实时检测算法的基音轨迹, 虚线为运用动态规划技术双向跟踪基音平滑处理的基音轨迹^[3]. 由上图可见该基音实时检测算法的效果与基音平滑处理的效果一样, 不仅能去除帧与帧之间基音周期不连续的



(a) 上图男声语音波形, 下图为对应的基音轨迹

(b) 上图女声语音波形, 下图为对应的基音轨迹

图2 语音信号的基音轨迹

跳跃点,而且能预测基音周期自然倍增现象的出现,提取可靠、准确的基音周期估计值。

大量的实验结果同样表明,该方法不仅具有基音平滑算法的准确性、可靠性,而且能在当前帧内实时地提取基音周期的估计值,减小基音平滑处理带来的额外编码延迟。

参考文献:

- [1] H. Ney. A dynamic programming technique for nonlinear smoothing [A]. in Proc. ICASSP. 81 [C], New York, 1981: 62- 65.
- [2] H. Ney. Dynamic programming technique for optional estimation of speech parameter contours [J]. IEEE Trans. on System, Man, and Cybernetics, March/ April 1983, SMC213: 208- 214.
- [3] 王都生, 铁满霞, 樊昌信. 一种实用的双向跟踪基音周期平滑算法 [J]. 电子学报, 1999, 27(10): 108~ 110.

作者简介:

王都生 1983 年获西安电子科技大学计算机应用专业工学学士学位, 1986 年、1999 年分别获西安电子科技大学通信与电子系统学科工学硕士、博士学位。现为西安电子科技大学副教授。研究兴趣为语音编码、数字信号处理。



铁满霞 1989 年获西安交通大学无线电技术专业工学学士学位, 1992 年获西安光机所计算机应用专业工学硕士学位。现为西安电子科技大学副教授。研究兴趣为语音编码、图象处理、通信网安全。

简 讯

一种新的雷达标校方法

全录贤

(西安卫星测控中心, 西安 710043)

摘 要: 本文研究了雷达的标校理论, 讨论了传统的望远镜标校方法, 提出了一种新的以电轴为基准的雷达标校方法。

1 引言

雷达标校, 包括标定和校准两部分工作。通过标定给出系统误差系数。校准是采取措施减小或消除系统误差。完成标校工作, 一般需要对系统误差进行数学描述, 即建立系统误差模型。通过误差模型的概念, 测控雷达的标校就变成建立适当的误差模型公式, 标定各个误差系数值, 采取某种措施来减小或消除系统误差。

传统上, 雷达的标校是以望远镜的光轴为基准来进行的, 对误差模型的建立也有很多的讨论。由于光轴是机械轴的中介, 而机械轴又是天线调整的几何轴心, 因而这种方法同光轴对机械轴的调整密切相关。理论上这种方法标校精度可以较高, 物理概念也清晰。但在实用上, 这种方法操作起来非常复杂, 例如光电标的架设要求很高, 望远镜不易准确对准目标, 望远镜的读数受人的视力和经验的影响, 标校室外作业, 受天气的限制等等。这使得其标校精度受到了限制, 而且也不能随时进行。

本文在讨论光学标校方法的基础上, 研究了以雷达的电轴为基准来直接进行标校的方法, 此种标校方法建立在天线是以机械轴为中心的对称结构的基础上, 实践证明, 它方便易行, 效率高, 且其标校精度高于光学标校。

2 光学标校

一般而言, 测量雷达有四种轴: 电轴、机械轴、光轴和编码器轴。

对于单脉冲雷达天线, 电轴就是差波束过零点的对称轴线。光轴是雷达天线座上安装的光学望远镜的物镜中心与目镜中心的连线。机械轴是与雷达天线方位轴与俯仰轴正交的轴。编码器轴由雷达轴角编码器决定。

雷达对目标进行跟踪测量时, 实际上是电轴在不断的追踪对准目标, 以编码器轴给出数据, 以机械轴和测站座标相联系。理想情况下此三个轴应是重合的。但实际上此三个轴是不可能完全重合的, 因而需要对此三个轴的不重合度进行标定。由于精确的机械轴难于确定, 光学标校中借助于望远镜光轴来进行标定。

一般情况下, 测量雷达需要修正的误差项主要有测量值的零位误差, 方位轴的不垂直度误差, 方位轴与俯仰轴的不正交度误差, 光轴和电轴的不匹配误差, 光轴和俯仰轴的不正交误差, 天线自身重力变形引起的重力下垂误差, 轴角编码器的偏心误差, 跟踪目标时的动态滞后误差, 以及电波在大气中传播时的折射误差。

其中方位轴与俯仰轴不正交度和轴角编码器安装偏心属于固定机械误差。前者在出厂前已测量好; 后者则在轴角编码器安装完成后用经纬仪和测角仪测量之, 直接在编码器中加以修正。两者以后一般情况下都不再测量。

方位轴不垂直度用合像水平仪测量, 运用最小二乘法归算出误差系数。动态滞后误差可由误差电压灵敏度直接进行修正。电波折射误差一般根据雷达跟踪时的气象参数, 由有关大气折射修正公式进行实时修正或事后修正。

其余零位误差, 光轴和电轴不匹配误差, 光轴和俯仰轴不正交误差, 重力下垂误差则需借助望远镜标定。此时, 需在标校塔上安装相应的光电标配合。电标架设要满足天线的远场条件, 并满足一定的高度; 光电标必须安装在同一平面, 该平面须垂直于电标与天线三轴原点的连线。光标共两个, 以电标为中心成奇对称配置, 相对于电标的位置同望远镜相对于天线三轴原点的位置。

(下转第 48 页)