

一种适合芯片技术的语音特征提取电路

王 阳[†]

(北京大学信息科学技术学院, 北京 100871)

摘要: 设计一种便于芯片技术实现、结构简单的语音特征提取电路. 它由带通滤波器组、整流电路和低通滤波器构成. 同时, 借此电路阐述一种充分利用器件特性设计电路的思想. 这种思想以完成功能为目的, 不以完成算法为目标. 据此建立的数学模型可能很复杂, 但电路结构简单, 便于芯片实现, 芯片资源利用率高. 对实际语音信号进行 SPICE 模拟结果初步表明, 此电路可以得到与线性系统类似的语音特征.

关键词: CMOS 集成电路; 模拟信号处理; 低功耗电路; 语音信号处理
EEACC: 2570D; 1205; 6130E

中图分类号: TN43 文献标识码: A 文章编号: 0253-4177(2006)S0-0374-04

1 引言

直接基于结构设计和优化电路而非器件, 可以充分发挥芯片技术的潜能, 提高芯片资源利用率^[1]. 如果设计电路使器件扩大工作范围, 可以充分利用器件电学特性, 进一步提高器件电学特性的利用率. 基于这种充分利用器件电学特性思想设计的电路, 以完成功能为目的, 不以完成算法为目标. 它的数学模型可能很复杂, 理论分析和计算机模拟分析会更困难, 但充分利用器件电学特性使电路结构简单, 便于芯片实现, 提高芯片资源利用率.

电子耳蜗的作用是将语音信号转换成能够有效代表语音特征的电学信号, 刺激耳蜗细胞, 修复生物耳蜗功能, 或输出到特征识别电路, 进行语音自动识别. 在没有真正弄懂生物耳蜗特征提取原理之前, 人们无法彻底知道如何设计类似生物耳蜗性能的特征提取电路, 只能在某些启示下进行各种探索. 根据耳蜗行波原理模型, 目前电子耳蜗实现方法主要有基于串联低通滤波器^[2~4]和并联带通滤波器^[5~8]两类. 本文以充分利用 CMOS 芯片物理结构设计的反馈环^[1]为基础, 设计一个充分利用器件特性、基于并联二阶带通滤波器的语音幅度均值特征提取电路. 它由带通滤波器组、半波整流电路和低通滤波器构成. 通过建立电路各基本部分数学模型, 研究扩大器件特性使用范围以及导致电路从线性变到非线性后产生的问题. 假设类似结构的线性系统提取的特征能够代表语音特征, 将它与线性系统比较, 评价提取特征的有效性.

2 幅度均值特征提取电路

从功能上理解耳蜗, 可以将其认为是一个并行带通滤波器组, 语音信号并行输入每一个带通滤波器. 因为人的听觉对相位不敏感, 可以有效地利用每个带通滤波器输出信号的幅度均值作为特征. 基于器件特性直接设计的幅度均值特征提取电路, 某一中心频率信号通道的简化电路如图 1 所示. 它由三部分组成: 二阶滤波器、半波整流电路和低通滤波器.

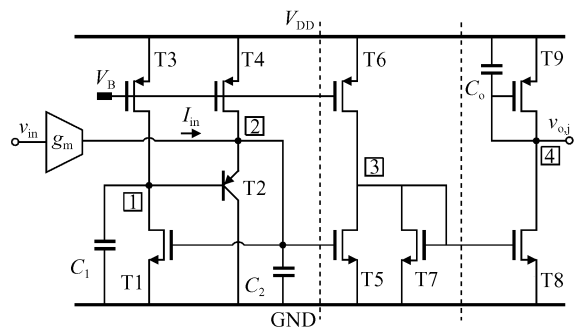


图 1 特征提取电路一个中心频率通道的简化电路

Fig. 1 Simplified schematic for a central frequency path in the circuit of feature extraction

为适合音频信号处理和减小功耗, MOS 管都设计在弱反型区工作. 因为整个电路的偏置电流很小, 器件失配对电路特性影响很大, 采用半波整流可以一定程度上减小器件失配对电路特性的影响.

对于 T1 和 T2 管构成的基于物理结构的二阶滤波器^[1], 选择电容端电压为状态变量, 根据电路建

[†] 通信作者. Email: ywang@pku.edu.cn
2005-10-11 收到, 2006-01-16 定稿

立系统状态方程：

$$\begin{cases} C_1 \frac{d\Delta V_1}{dt} = I_{D0} (1 - e^{\frac{q\Delta V_2}{nkT}}) - \frac{I_{E0}}{1+\beta} (1 - e^{\frac{q}{kT}(\Delta V_2 - \Delta V_1)}) \\ C_2 \frac{d\Delta V_2}{dt} = I_{E0} (1 - e^{\frac{q}{kT}(\Delta V_2 - \Delta V_1)}) + \Delta I_{in} \end{cases} \quad (1)$$

其中 $\Delta V_1(t) = V_1(t) - V_{10}$; $\Delta V_2(t) = V_2(t) - V_{20}$; $I_{D0} = I_{Ds} \exp(qV_{20}/nkT)$; $I_{E0} = I_{Es} \exp(q(V_{20} - V_{10})/kT)$; β 是双极管的电流增益; n 是 MOS 管的亚阈值斜率因子. 这种充分利用器件特性设计的电路, 具有很大的非线性, 它追求的不是实现经典的线性算法而是实现功能. 这样虽然能使结构变得简单, 但却使数学分析和计算机模拟变得更加困难. 因此, 它只适合芯片实现, 不适合算法实现.

当外部输入为零时, 系统状态点的演变情况由系统动力学特性决定. 在平衡点 $(\Delta V_{1c}, \Delta V_{2c})$ 处, $d\Delta V_{1c}/dt = 0$, $d\Delta V_{2c}/dt = 0$, 状态方程(1)变为非

线性方程组, 由此解出的平衡点位于零点.

(1)式在平衡点附近的小信号线性近似结果为:

$$\begin{cases} C_1 \frac{d\Delta V_1}{dt} = -g_{m1} \Delta V_2 + \frac{1}{1+\beta} g_{m2} (\Delta V_2 - \Delta V_1) \\ C_2 \frac{d\Delta V_2}{dt} = -g_{m2} (\Delta V_2 - \Delta V_1) + \Delta I_{in} \end{cases} \quad (2)$$

其中 $g_{m1} = qI_{D0}/nkT$; $g_{m2} = qI_{E0}/kT$. 由于系数矩阵的特征值实部都不为零, 所以平衡点是双曲平衡点. 根据 Hartman 定理, 非线性系统双曲平衡点邻域的相图与线性化原点邻域的相图从定性上讲是等价的, 它们有类似的不稳定流形和不稳定流形. (1)式所描述系统的大信号和小信号相图如图 2 所示, 可见它们有类似的稳定性, 大信号非线性应用时不会引起稳定性的变化.

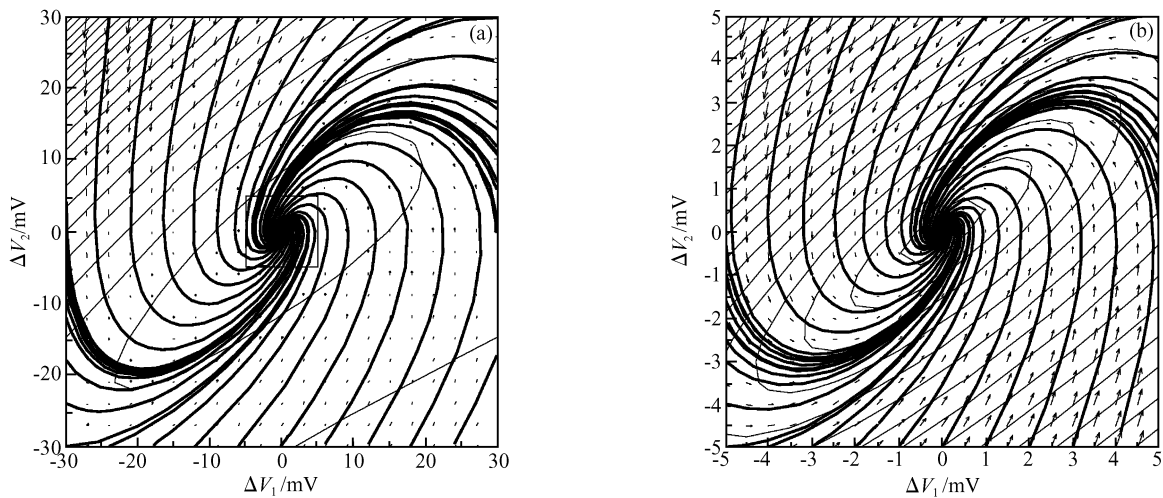


图 2 非线性二阶滤波器的平衡点附近相图 (a)大信号相图; (b)小信号相图

Fig.2 Phase diagrams of the second order nonlinear filter near equilibrium point (a) Large signal; (b) Small signal

当加外部输入信号后, 电路的各个输出随信号频率特性不同而变化. 直接从非线性系统方程(1)出发分析频率特性很不方便, 因此先在小信号线性近似的情况下, 用经典的线性时不变系统理论分析频率特性. 对于大信号非线性情况, 将其认为是小信号加畸变的结果. 根据(2)式小信号线性化近似结果, 可得:

$$\begin{aligned} [s^2 C_1 C_2 + s g_{m2} (C_1 + \frac{C_2}{1+\beta}) + g_{m1} g_{m2}] \Delta V_2 \\ = (s C_1 + \frac{g_{m2}}{1+\beta}) \Delta I_{in} \end{aligned} \quad (3)$$

可见输入电流与节点 2 电压之间表现出带通滤波特性, 并可由此得到中心频率和质量因子.

当信号相对于偏置量增大后, 实际输出偏离线性化近似结果. 如果将偏离量看成是非线性引起的畸变, 这种畸变将使输出产生更多的频率分量. 但最终特征量是幅度均值, 这种非线性畸变对幅度均值影响多大以及线性系统得到的幅度均值表征语音信号的特征程度如何, 因为没有严格理论证明线性系统得到的幅度均值能够最好地表示语音特征, 所以畸变后的幅度均值是增强了特征还是减弱了特征尚需进一步理论和实验的研究.

半波整流由电压-电流转换电路和电流型半波整流电路完成. 电压-电流转换电路由 T5、T6 管构成. T7 管实现电流半波整流, T8 管将其输出给后续低通滤波器. T7 管电流为:

$$I_{D7} = \begin{cases} 0, & I_{SD6} \leq I_{D05} e^{\frac{q\Delta V_2}{nkT}} \\ I_{SD6} - I_{D05} e^{\frac{q\Delta V_2}{nkT}}, & I_{SD6} > I_{D05} e^{\frac{q\Delta V_2}{nkT}} \end{cases} \quad (4)$$

其中 I_{SD6} 是 T6 管的源漏电流; I_{D05} 是信号为零时 T5 管的漏极电流.

低通滤波由 T9 管和电容 C_o 实现, 输出信号的变化量满足微分方程:

$$C_o \frac{d\Delta V_{out}}{dt} + I_{D09} e^{\frac{q\Delta V_{out}}{nkT}} = I_{D7} \quad (5)$$

小信号线性化近似后, 上式变为:

$$C_o \frac{d\Delta V_{out}}{dt} + \frac{q}{nkT} I_{D09} \Delta V_{out} = I_{D7} - I_{D09} \quad (6)$$

截止频率为: $g_{mo}/2\pi C_o$, 其中 $g_{mo} = qI_{D09}/nkT$, I_{D09} 是输入信号为零时 T9 管漏极电流. 理想情况下, $I_{D09} = I_{SD6} - I_{D05}$. 低频滤波特性一方面由滤波电容 C_o 决定, 另一方面由 I_{D09} 决定, 因此合理控制 I_{D09} 变得很重要.

3 SPICE 模拟分析

模拟分析采用 16 路带通滤波通道, 在 Mel 频率坐标上等间隔中心处选择中心频率, 频率范围在 200~4000Hz 之间, MOS 管采用 BSIM3v3.1 模型, 用数字“二”的语音作输入信号, 如图 3 所示. 由于这个电路设计是建立在充分利用器件基本特性基础之上, 追求的目标是实现有效的特征提取, 而不是精确地实现滤波算法等, 所以只对最终输出能否有效代表特征进行评价. 假设现在常用线性系统得到的是有效特征, 将图 1 所示非线性系统得到的结果与线性系统进行比较, 观察非线性系统能否保持类似特性.

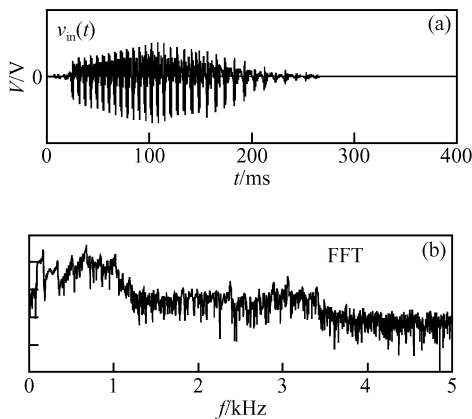


图 3 “二”的语音信号 (a) 电压信号; (b) FFT 频谱
Fig.3 Speech signal of “2” in Chinese (a) Voltage signal; (b) FFT spectrum

评价特征所用线性系统采用理想跨导器构成二阶带通滤波器, 与非线性带通滤波器具有相同的小

信号频率响应特性. 线性系统的输出经过理想跨导转换成电流输入到图 1 所示的半波整流电路. 一组非线性与线性系统输出的典型结果如图 4(a)和(b)所示, 其中输出幅度均值是线性坐标. 从结果可以看到幅度均值的变化趋势非线性系统与线性系统基本相同. 这种比较还很初步和简单, 但自我作故, 无所究章, 如何更科学的评价特征的有效性尚待研究.

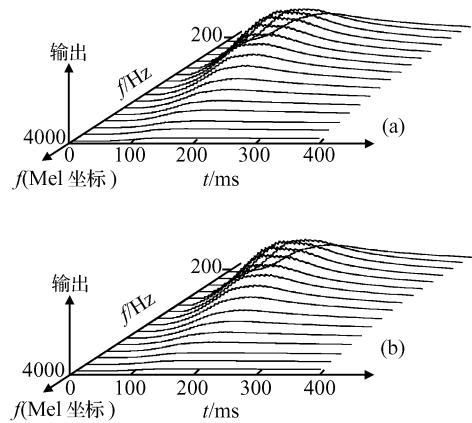


图 4 非线性系统与线性系统输出结果比较, 输出是线性坐标 (a) 非线性滤波器; (b) 线性滤波器
Fig.4 Comparison of output results for nonlinear and linear systems, linear scale for outputs (a) Nonlinear filter; (b) Linear filter

与输入信号的 FFT 频谱(图 3(b))比较可以看到, 二阶带通滤波器实现的特征提取电路频率分辨率较低, 这主要是由于二阶带通滤波器的半功率带宽较大所致, 只通过增加滤波器数目不能有效提高分辨率. 虽然它可以用于耳蜗系统, 但作为语音识别的特征或高性能电子耳蜗, 频率分辨率还不够. 解决分辨率问题可以采用进一步增加滤波器的阶数和提高滤波器 Q 值的方法.

4 结论

充分利用 CMOS 芯片物理结构和器件电学特性可以设计出结构简单、便于硬件实现、芯片资源利用率高的电路. 但是, 充分利用器件电学特性, 以完成功能为目的, 不以完成算法为目标. 但使电路从线性扩展到非线性工作区导致数学模型复杂, 理论分析和计算机模拟分析更加困难. 对于根据这种思想设计的基于二阶带通滤波器组的语音幅度均值特征提取电路, 用实际语音信号进行 SPICE 分析结果初步表明, 它的最终输出与线性系统有类似的语音特征.

参考文献

- [1] Wang Y. A feedback loop structure based on CMOS technology and its natures. The Proceedings of the 13th National Conference on Integrated Circuits and Silicon Materials, 2003:357(in Chinese)[王阳.一种基于 CMOS 技术的反馈环结构及其特性.第十三届全国半导体集成电路与硅材料学术会议论文集,2003:357]
- [2] Watts L, Kerns D A, Lyon R F, et al. Improved implementation of the silicon cochlea. IEEE J Solid-State Circuits, 1992, 27(5):692
- [3] Furth P M, Andreou A G. A design framework for low power analog filter banks. IEEE Trans Circuits Syst- I, 1995, 42(11):966
- [4] Fragniere E, Van Schaik A, Vittoz E A. Design of an analogue VLSI model of an active cochlea. Analog Integrated Circuits and Signal Processing, 1997, 12(3):19
- [5] Wilson B S, Finley C C, Lawson D T, et al. Speech processors for cochlear prostheses. Proc IEEE, 1988, 76(9):1143
- [6] Unnikrishnan K P, Hopfield J J, Tank D W. Connected-digit speaker-dependent speech recognition using a neural network with time-delayed connections. IEEE Trans Signal processing, 1991, 39(3):698
- [7] Wang Y. Continuous-time filter banks on subthreshold CMOS for sound analyses. 3rd International Conference on ASIC Proceedings, 1998:151
- [8] Georgiou J, Toumazou C. A 126- μ W cochlear chip for a totally implantable system. IEEE J Solid-State Circuits, 2005, 40(2):430

A Speech Feature Extraction Circuit Suitable for Chip Technology

Wang Yang[†]

(School of Electronics and Computer Science, Peking University, Beijing 100871, China)

Abstract: A speech feature extraction circuit that is simple and suitable for chip technology is designed. It consists of a filter bank, rectifying circuit, and low pass filters. Meanwhile, an idea of the device characteristics efficient for designing circuits is illustrated with this circuit. The idea takes implementation of the functions rather than the algorithms as the purpose. Although the mathematical models are complicated, the circuits designed according to this idea are simple, suitable for chip implementation, and more efficient for the limited resources on chip. The results show that the circuit can obtain the features similar to the linear system through the SPICE simulation for practical speech signals.

Key words: CMOS integrated circuit; analog signal processing; micropower circuit; speech signal processing

EEACC: 2570D; 1205; 6130E

Article ID: 0253-4177(2006)S0-0374-04

[†] Corresponding author. Email: ywang@pku.edu.cn

Received 11 October 2005, revised manuscript received 16 January 2006