

分段式高性能近似加法器设计

李凯磊, 杨文强, 王伦耀*

(宁波大学 信息科学与工程学院, 浙江 宁波 315211)

摘要: 利用近似加法器输出误差与输出位权重有关的特点, 提出了一种分段式近似加法器结构. 该近似加法器由三部分组成, 权重高部分由精确加法器构成, 权重低部分的输出被置为全 1, 而中间部分由二输入 OR 门代替全加器. 提出的近似加法器与精确行波进位加法器相比, 在面积、功耗、功耗延迟积和面积延迟积分别改善了约 43%、57%、85%和 81%, 与已有的近似加法器相比, 总体性能也有明显提升.

关键词: 近似加法器; 行波加法器; 近似计算; 低功耗

中图分类号: TP331

文献标志码: A

文章编号: 1001-5132 (2022) 06-0009-06

在数字电路设计中, 逻辑正确是设计的首要目标. 但在实际应用中, 如涉及图像、音频处理等多媒体应用方面, 因人类感知能力限制, 即使存在一定范围内的错误输出, 也不影响人们的视听效果. 实际应用中的容错特性为逻辑电路的优化提供了一条新途径. 利用近似计算技术实现电路逻辑优化, 其本质就是通过降低电路的输出精度来实现电路简化, 进而达到电路的功耗、面积或者时延等性能指标的优化.

在集成电路设计中, 加法器是一种被广泛应用的电路单元, 加法器性能与整个系统性能关系密切. 高性能加法器设计一直是学术界和工业界高度关注的问题^[1-2]. 在误差允许的情况下, 引入近似加法器是提高加法器性能的一种有益尝试. 考虑到近似加法器的输出误差与发生错误输出位的权重有关, 权重到高的输出位发生错误将引起更大的误差, 并且输出误差的大小也是近似加法器设计过程中一个关键指标, 为此本文提出了一种三段式加法器设计, 该设计基本思想是在低权重输出位用简单的逻辑门, 甚至常数来代替全加器, 而在高权重部分则采用精确加法器来实现, 从而在误差大小与电路性能之间实现折中.

1 相关工作

1.1 研究现状

已有的近似加法器在结构上大致分为二段式和多段式两种类型^[3].

二段式近似加法器包含两部分, 即由精确加法器构成 MSB(More Significant Bits)部分, 以及由近似加法器构成的 LSB(Less Significant Bits)部分, 其中, LSB 部分一般采用无进位近似加法器. 此类近似加法器的每一位都单独进行运算, 不产生或接收进位信号, 因而可以减少因进位产生的延时, 进而提高电路的运算速度.

二段式近似加法器的这种结构可以在增加有限输出误差的情况下, 实现电路面积、延迟以及功耗等优化. 现有的低位或门加法器(Lower Part OR Adder, LOA)^[4]、容错加法器 I(Error Tolerant Adder-I, ETA-I)^[5]等均为二段式近似加法器.

其中, ETA-I 加法器的精确部分由行波进位加法器实现, 近似部分每一位求和运算全部使用 XOR 模块代替. 该模块在第 i 位输入 A_i 和 B_i 分别在“00”“01”或“10”的时候进行正常 XOR 运算; 而在“11”的时候, 从第 i 位开始至最低位, 求和的输

收稿日期: 2022-01-06.

宁波大学学报(理工版)网址: <http://journalg.nbu.edu.cn/>

基金项目: 国家自然科学基金(61871242); 浙江省自然科学基金(LY19F040004).

第一作者: 李凯磊(1997-), 男, 浙江余姚人, 在读硕士研究生, 主要研究方向: 逻辑综合与优化. E-mail: 1324992817@qq.com

*通信作者: 王伦耀(1972-), 男, 浙江宁波人, 教授, 主要研究方向: 逻辑综合与优化. E-mail: wanglunyao@nbu.edu.cn

出结果全部被设置为“1”. 此类加法器的输出误差比较小, 但因其高位部分由传统的精确加法器构成, 且通过 XOR 门以及控制模块组成低位, 因此相对于某些近似加法器电路, 其电路面积开销较大.

现有多段式近似加法器包括精度可配置加法器 II (Accuracy Configurable Approximate Adder-II, ACA-II)^[6]、容错加法器 II (Error Tolerant Adder-II, ETA-II)^[7]、预测进位选择加法器 (Speculative Carry Select Addition, SCSA)^[8] 和通用精度可配置加法器 (Generic Accuracy Configurable Adder, GeAr)^[9] 等. 多段式加法器一般由多个并行子加法器模块组成, 每个子加法器模块分为 MSB 部分与 LSB 部分. 其中 MSB 部分用于输出一部分求和结果, LSB 部分用于预测进位信号, 并传递给 MSB 部分, 不用于求和计算. 最后所有子加法器 MSB 部分的求和结果组合起来即为此类近似加法器的最终求和输出. 此类多段式加法器虽然可以显著降低延迟, 但是因其由多个子加法器模块构成, 电路面积相较于其他近似加法器增加很多.

ETA-II 加法器结构如图 1 所示, n 位加法器被分割成 m 个子加法器模块, 每个模块中包含 1 个 k 位行波进位加法器 (Ripple Carry Adder, RCA) 模块, 以及 1 个 k 位超前进位加法器 (Carry Lookahead Adder, CLA) 模块, 分别用于 2 个加数对应字段求和以及进位信号预测.

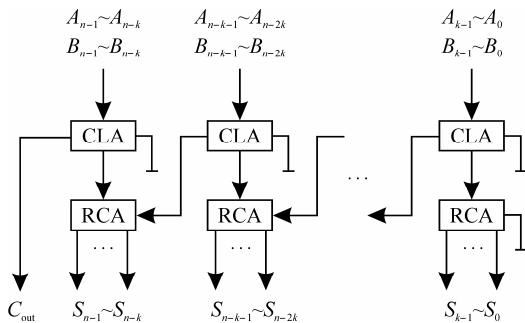


图 1 ETA-II 加法器

从图 1 可知, ETA-II 多段式加法器第 h 个子加法模块的进位仅仅与当前 k 位输入和第 $(h-1)$ 个子加法模块的 k 位输入有关, $1 \leq h \leq (m-1)$, 因此该加法器可以将进位链深度由原来的 n 缩短到 $2k$, 进而显著减小电路时延. 但同时因进位链被截断, 使得在某些输入组合下, ETA-II 会产生很大的误差.

以图 2 所示 12 位加法器为例, 该加法器包含 3 个子加法模块, 本文称之为 ETA-II-3 加法器. 当输入 2 个 12 位的加数为 $A[11:0] = "000011111111"$ (255) 和 $B[11:0] = "000000000001"$ (1) 时, 由于处于高位子加法模块的进位来自于次高位子加法电路的进位输出, 在图 2 所示的输入组合时, 次高位子加法电路的进位输出为 0, 导致该 12 位加法器最终的输出为 0, 而不是 256, 产生很大误差.

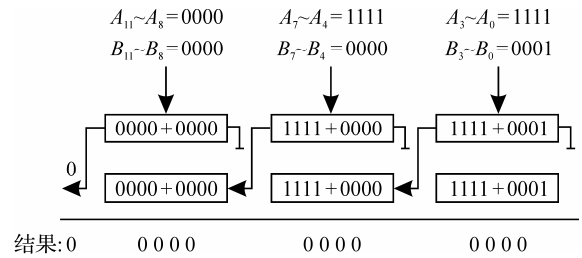


图 2 ETA-II-3 产生误差

1.2 计算误差评价指标

近似加法器的计算误差通常采用误差距离 (Error Distance, ED), 以及与误差距离紧密相关的平均误差距离 (Mean Error Distance, MED)、相对误差距离 (Relative Error Distance, RED)、平均相对误差距离 (Mean Relative Error Distance, MRED) 和归一化误差距离 (Normalized Mean Error Distance, NMED) 等指标来衡量^[10].

误差距离 (E_{ED}) 定义为近似加法器所获得的值与正确结果之间差值的绝对值. 式(1)中 R_c 和 R_e 分别表示正确值和近似值, 且均用十进制表示:

$$E_{ED} = |R_c - R_e|. \quad (1)$$

平均误差距离 (E_{MED}) 为 α 个误差距离取平均值:

$$E_{MED} = \frac{1}{\alpha} \sum_{i=0}^{\alpha-1} E_{ED,i}. \quad (2)$$

相对误差距离 (E_{RED}) 是误差距离与正确值比:

$$E_{RED} = \frac{E_{ED}}{R_c}. \quad (3)$$

平均相对误差距离 (E_{MRED}) 为 α 个相对误差距离取平均值:

$$E_{MRED} = \frac{1}{\alpha} \sum_{i=0}^{\alpha-1} E_{RED,i}. \quad (4)$$

归一化误差距离是平均误差距离的归一化:

$$E_{NMED} = \frac{E_{MED}}{2^0 + 2^1 + \dots + 2^{\alpha-1}}. \quad (5)$$

加法器精度(Accuracy, ACC)表示特定输入的输出“正确”程度, 其值范围为0%~100%, 定义为:

$$A_{ACC} = \left(1 - \frac{E_{ED}}{R_c}\right) \times 100\%. \quad (6)$$

最低可接受精度(Minimum Acceptable Accuracy, MAA)表示近似加法器输出结果精度最低的百分值, 而可接受结果概率(Acceptance Probability, AP)是指加法器精度高于最小可接受精度的概率, $0 \leq A_{AP} \leq 1$:

$$A_{AP} = P(A_{ACC} > M_{MAA}). \quad (7)$$

2 三段式高性能非精确加法器

2.1 三段式低功耗非精确加法器介绍

本文提出的三段式 n 位高性能非精确加法器结构如图3所示, $n=4z, z \in \{1, 2, 3, 4, \dots\}$. 该近似加法器被分割成三部分. 第一部分为加法器低位部分, 位宽为 $n/4$, 这部分每一位的输出结果被设置为常数“1”, 也称为 CON 块; 第二部分为加法器中间部分, 位宽为 $n/4$, 用一个二输入逻辑门代替一位全加器, 这部分电路也称 TP 块. 剩余部分为精确计算部分, 由 RCA 构成.

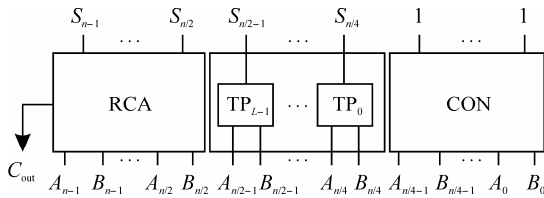


图3 三段式高性能非精确加法器结构图

2.1.1 CON 块

CON 块为常数块, 即无论 $A[n/4-1:0]$ 和 $B[n/4-1:0]$ 输入什么值, 输出结果 $S[n/4-1:0]$ 将全部被置为常数, 这样将使延迟、面积和功耗得到大幅度降低. 由于平均误差距离是衡量近似加法器的一个关键指标, CON 块的常数输出设置应该尽可能引入较小的平均误差.

设 A 、 B 为 2 个加数, A 、 B 为整数, 且 $0 \leq A \leq K-1, 0 \leq B \leq K-1$. 则 A 、 B 可以产生 K^2 种输入组合. 设 S_i 为其中一种输入组合的和, $0 \leq i \leq (K^2-1)$, 则所有输入情况下加法器输出的求和结果为:

$$S = \left\{ \frac{[0 + (K-1)]K}{2} + \left[\frac{(K-1)K}{2} + (K-1)K \right] \right\} \times$$

$$\frac{K}{2} = (K-1)K^2, \quad (8)$$

因此, 其求和的平均值为:

$$S_{avg} = S / K^2 = K-1. \quad (9)$$

显然, 对于位宽为 $n/4$ 的 CON 块, 为了使 CON 块对应的平均误差距离最小, CON 块的输出应该为 $2^{n/4}-1$, 即置 CON 块的所有输出为“1”.

2.1.2 TP 块

TP 块输出结果 $S[n/2-1:n/4]$ 由 $A[n/2-1:n/4]$ 和 $B[n/2-1:n/4]$ 的每一位单独进行基本二输入逻辑运算所得, 且每一位不产生也不接收任何进位信号, 从而减少电路面积和时延.

从表1可知, XOR 门的输出与加法器的求和输出一致, 但没有进位信号; OR 门的 4 个输出中有 3 个与加法器的输出一致, 而 AND 门只有 1 个输出与加法器的输出一样. 因此, 在 TP 块中不考虑 AND 门.

表1 TP 块逻辑门选择与半加器结果比较

输入		加法器		XOR 门	OR 门	AND 门
A_i	B_i	和	进位			
0	0	0	0	0	0	0
0	1	1	0	1	1	0
1	0	1	0	1	1	0
1	1	0	1	0	1	1

相比 OR 门, 虽然在 4 个输入组合中, XOR 门的输出均与加法器和的输出一样, 但由于在 TP 块中各位之间没有进位, 因此在 2 个加数的某一位输入均为“1”时, XOR 门的输出实际上是“0”, 而 OR 门的输出为“1”, 反而 OR 门引起的误差更小. 另外, 对于二输入 OR 门而言, 只有其中一个输入为“0”时, 另一个的输入变化才能引起输出变化, 而二输入 XOR 门的任何输入变化都会反映到输入端, 因此相比 OR 门, XOR 门对输入的变化更敏感, 其动态功耗也将更大. 此外, XOR 的电路面积要比 OR 门大. 在本文中, 将同时给出 OR 门和 XOR 替代全加器后的电路性能比较结果.

2.1.3 RCA 块

三段加法器的高位部分由精确加法器组成. 对于一个 n 位二进制, 其对应的十进制可表示为:

$$V = \sum_{i=0}^{n-1} (B_i \times 2^i), B_i \in \{0, 1\},$$

从上式可知, 由于高位权重大于低位权重, 因此同样某一位上出错, 高位错误引起的误差会更大, 这也就是 TSA 加法器在高位采用精确加法器的原因. 本文设计中, 高位部分采用行波加法器, 此类加法器结构简单, 但进位链长. 由于本设计中将进位链进行了截断, 长度只有原来的一半, 因此可有效减少整个电路的时延.

2.2 三段式非精确加法器误差分析及改进

由图 3 结构可知, 该近似加法器产生误差的情况均取决于 TP 块和 CON 块. 相比与 CON 块, TP 块在图 3 的结构中处于高位段. 因此 TP 块和 CON 块引起误差的大小主要由 TP 块决定.

从表 1 可知, 在二进制加法中, 某一位的 2 个加数的 4 种输入组合中, 当 TP 块中的二输入逻辑门采用 XOR 门时, 求和输出均正确, 但是在输入组合为“11”时, 没有产生进位. 当 TP 块中的二输入逻辑门采用 OR 门时, 在输入组合为“11”时, 求和输出错误, 同时没有产生进位, 其他 3 种输入组合均正确. 从上面分析不难看出, TP 块产生的求和输出结果小于正确值. 为减少 TP 块的误差距离, 本文将 TP 进行了如下改进.

假设 TP 块长度为 L , 则将第 $\lfloor L/2-1 \rfloor$ 位的二输入逻辑门输出用常量“1”来代替. 同时, 考虑到 TP 块最高位为“1”时, 肯定会产生进位输出, 因此将 TP 块结构改进为图 4 所示结构.

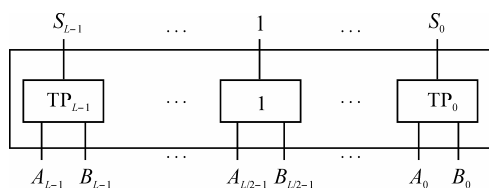


图 4 改进后 TP 块结构图

表 2 为在位宽为 16 位情况下, 图 4 结构的 TP 块与图 3 结构的 TP 块在随机输入情况下与正确值的错误误差比较结果.

表 2 修改后 16 位 TP 模块加法精度比较

项目	MED	NMED	MRED	功耗/ μW	面积/ μm^2
TP _{OR}	64.25	0.252 0	0.213 3	20.69	2 294.33
TP _{XOR}	128.50	0.503 9	0.426 5	47.29	2 335.06
TP _{OR-1}	63.20	0.247 8	0.210 6	18.03	2 007.54
TP _{XOR-1}	109.75	0.430 4	0.399 6	41.48	2 043.18

表 2 中, TP_{OR} 和 TP_{XOR} 分别表示 TP 块中二输

入逻辑门采用 OR 门和 XOR 门后电路的各性能参数, TP_{OR-1} 和 TP_{XOR-1} 分别表示采用图 4 结构后的电路性能参数. 并且表 2 中, 各误差参数是在随机产生 10^5 种输入组合, 利用 MATLAB 模拟得到. 面积和功耗参数利用 Design Compiler 在 fast_1v32c0 工艺库下得到.

通过表 2 可知, 在 8 位宽下, TP 块采用 OR 门所产生的计算误差相较于采用 XOR 门要小, 且在电路功耗和电路面积方面的表现也比采用 XOR 门要好. 采用图 4 结构的 TP_{OR-1} 与仅由 OR 门构成的 TP_{OR} 相比, 在计算精度和减小电路开销方面均有提升.

结合表 2 的结果, 本文提出的近似加法器的结构如图 5 所示, 并称该电路为 TSA 加法器. 其中, 在 TP 块中采用二输入 OR 门.

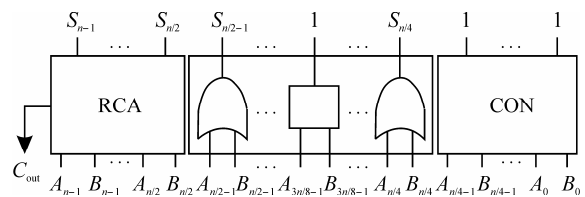


图 5 TSA 加法器结构图

3 实验结果

本节将对文中提出的近似加法器与几种经典的近似加法器在面积、延迟、功耗和精度等方面的特性进行比较. 为获得加法器电路的面积、功耗和时延数据, 先将近似加法器用 Verilog HDL 语言描述, 然后用 Design Compiler 进行综合得到, 其中综合过程中用到的工艺库为 fast_1v32c0. 表 3 即为实验结果.

表 3 各 16 位加法器电路性能比较

加法器类型	面积/ μm^2	延迟/ns	功耗/ μW	PDP/ $(\mu\text{W}\cdot\text{ns})$	ADP/ $(\mu\text{m}^2\cdot\text{ns})$
RCA	15 694	5.13	547	2 806	80 509
LOA	9 931	1.76	248	437	17 478
ETA-I	13 674	1.76	340	597	24 067
ETA-II-2	16 179	2.99	550	1 645	48 376
ETCA	10 399	1.76	259	455	18 303
TSA	8 870	1.76	236	415	15 612

表 3 中, 近似加法器 LOA、ETA-I、ETA-II-2、ETCA^[11]和 TSA 的加数位宽均为 16 位, 且高 8 位

均采用 RCA 精确加法器. 从表中可以看出, 本文所提出的 TSA 近似加法器总体性能较好. 由于 TSA 中间部分采用部分精确电路, 通过 OR 门代替一位全加器, 低位部分每一位输出结果都被设置为常数 1, 这种设计有利于 TSA 在面积、功耗和时延上的优化. 相较于 RCA 面积、功耗、功耗延迟积(Power Delay Product, PDP)和面积延迟积(Area Delay Product, ADP)分别下降约 43%、57%、85% 和 81%. 与 ETCA 相比, TSA 加法器在面积、功耗和面积功耗积方面分别有大约 15%、9%、22% 的改善, 相对于 LOA、ETA-I、ETA-II-2 各方面都具有不同程度的提高. 由于上述电路均采用进位链截断技术, 进位链长度相仿, 因此各电路的时延均接近. 而 ETA-II-2 由于将求和运算与进位信号预测分为两部分, 导致时延比较长.

计算误差是衡量近似加法器的一个关键指标. 本文中各近似加法器的计算误差是利用 MATLAB 对各近似加法器进行建模, 并输入 10^5 个随机数进行仿真获得. 表 4 即为各近似加法器计算误差仿真结果.

表 4 各 16 位加法器计算误差比较

加法器类型	MED	NMED	MRED
LOA	63.47	0.000 968	0.001 920
ETA-I	50.83	0.000 776	0.001 561
ETA-II-2	126.51	0.001 930	0.003 884
ETCA	51.70	0.000 789	0.001 565
TSA	61.07	0.000 932	0.001 291

不难看出, TSA 近似加法器因 TP 块采用 OR 门, 低位部分直接被置为“1”, 其计算误差与 ETCA 和 ETA-I 相比略有增加, 但相较于 ETA-II-2 和 LOA 在平均误差距离、归一化误差距离以及平均相对误差距离均有改进, 其中与 ETA-II-2 相比, 上述指标分别提高了约 52%、52% 和 67%.

对不同位宽的 TSA 加法器和 MAA 下 AP 的仿真结果见表 5. 对于不同 MAA(MAA<100%)值, 位宽越大, 其 AP 值越趋近于 100%, 因此本文所提出的加法器在位宽较大的加法计算中也具有极大优势.

4 总结

由于加法器的计算误差大小与错误发生的位

表 5 不同位宽和 MAA 下 TSA 加法器的 AP 值 %

MAA	AP			
	8 位	12 位	16 位	32 位
100	14.883	5.257	1.906	0.041
99	40.886	76.295	98.151	100
98	59.434	93.338	99.599	100
97	69.420	97.211	99.805	100
96	78.094	98.337	99.892	100
95	84.943	98.941	99.938	100
94	90.055	99.306	99.960	100
93	92.319	99.536	99.964	100
92	94.414	99.630	99.968	100
91	95.482	99.730	99.974	100
90	96.460	99.770	99.992	100

置关系密切, 相比于权重大的输出位发生错误, 权重小的输出位错误引起的误差要小很多. 针对加法器此特点, 本文三段式近似加法器将加法器分成三段, 其中权重高的段用精确加法器实现, 达到降低误差目的, 中间和低位部分采用 OR 门和常数, 且各位之间不存在进位关系. 这种设计显著降低了电路面积和功耗, 同时也对计算误差与电路性能改善进行了折中. 与传统精确加法器行波进位加法器相比, 该加法器的面积、功耗、功耗延迟积和面积延迟积分别下降约 43%、57%、85% 和 81%. 与现有分段式近似加法器相比, 总体性能有明显改善.

参考文献:

- [1] Wang Y, Dong J, Liu Y X, et al. RMLIM: A runtime machine learning based identification model for approximate computing on data flow graphs[J]. IEEE Transactions on Sustainable Computing, 2022, 7(1): 201-210.
- [2] Balasubramanian P, Dang C, Maskell D L. Approximate quasi-delay-insensitive asynchronous adders: Design and analysis[C]//2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS). Boston, MA, USA. IEEE, 2017:1196-1199.
- [3] Kim Y. An accuracy enhanced error tolerant adder with carry prediction for approximate computing[J]. IEEE Transactions on Smart Processing & Computing, 2019, 8(4):324-330.
- [4] Mahdiani H R, Ahmadi A, Fakhraie S M, et al.

- Bio-inspired imprecise computational blocks for efficient VLSI implementation of soft-computing applications[J]. IEEE Transactions on Circuits and Systems Part I: Regular Papers, 2010, 57(4):850-862.
- [5] Zhu N, Goh W L, Zhang W, et al. Design of low-power high-speed truncation-error-tolerant adder and its application in digital signal processing[J]. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2010, 18(8):1225-1229.
- [6] Kahng A B, Kang S. Accuracy-configurable adder for approximate arithmetic designs[C]//DAC '12: Proceedings of 49th Annual Design Automation Conference, 2012: 820-825.
- [7] Ning Z, Goh W L, Yeo K S. An enhanced low-power high-speed adder for error-tolerant application[C]// Proceedings of the 12th International Symposium on Integrated Circuits. Singapore. IEEE, 2009:69-72.
- [8] Du K, Varman P, Mohanram K. High performance reliable variable latency carry select addition[C]//2012 Design, Automation & Test in Europe Conference & Exhibition (DATE). Dresden, Germany. IEEE, 2012: 1257-1262.
- [9] Shafique M, Ahmad W, Hafiz R, et al. A low latency generic accuracy configurable adder[C]//DAC '15: Proceedings of the 52nd Annual Design Automation Conference. 2015:1-6.
- [10] Jiang H L, Liu C, Liu L B, et al. A review, classification, and comparative evaluation of approximate arithmetic circuits[J]. ACM Journal on Emerging Technologies in Computing Systems, 2017, 13(4):1-34.
- [11] Seo H, Yang Y S, Kim Y. An energy-efficient imprecise adder with a lower-part constant approximation[C]//2020 International SoC Design Conference (ISOCC). Yeosu, South Korea. IEEE, 2020:143-144.

Design of segmenting high performance approximate adder

LI Kailei, YANG Wenqiang, WANG Lunyao*

(Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China)

Abstract: Taking advantage of the feature that the error distance of an approximate adder is closely related to the weight of the corresponding bit with error output, a segmented approximate adder structure is proposed in this paper. The approximate adder includes three parts: the part with the highest weight being composed of an accurate adder, the part with the lowest weight having each output set to 1, and the middle part consisting of two-input OR gates instead of the full adders. Compared with the ripple carry adder, the proposed approximate adder attains significant improvements in area, power consumption, power delay product, and area delay product by about 43%, 57%, 85%, and 81% respectively. Compared to the existing approximate adders, the overall performance is also improved significantly.

Key words: approximate adder; ripple carry adder; approximate calculation; low power

(责任编辑 章践立)