1. 问题分析

构造一个 3*3 的位宽 16bit 卷积核,并利用该卷积核完成与 7*7 矩阵的卷积运算,数据位宽 16bit 补码数,乘法结果位宽 32bit 补码数,降低精度截断成 16 位输出。

卷积的基本过程如下:

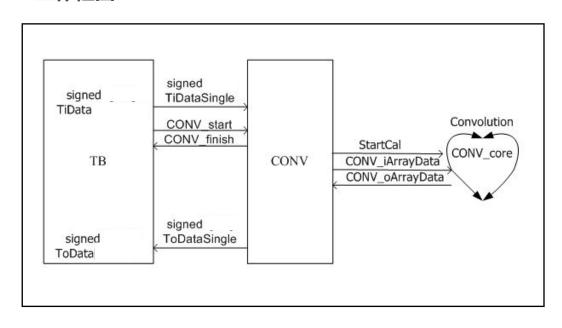
Value 形式的卷积: 33 卷积核的边缘对准 77 矩阵的边缘,从输入矩阵 pic 的左上角开始,以步长为 1 向右移动,触边后向下移动,按此路径到达右下角结束。卷积核 convolution core 的每个数与 pic 矩阵对应位置数据相乘累加得出 1个 conv_multi 结果,如此往复作业,最终输出卷积后的 5*5 矩阵 conv_oData。

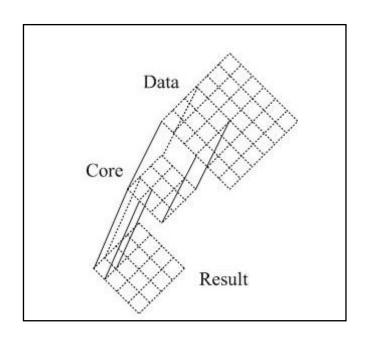
前仿流程:将上述过程用 verilog 代码描述,得到 CONV.v 文件,后用 testbench 文件 CONV_tb.v 给予测试信号,编译并使用 verdi 前仿软件进行仿真。

后仿:使用 DC 软件仿真,得到 CONV_post.v 以及时序约束文件 CONV.sdc

综合: 使用 Innovus 软件布局布线。

2. 工作框图





3. 数据选取

系统内置 3*3 的 16bit 补码数的卷积核

FD01	01CD	FEFC
FFC4	FF30	FDB4
FEDD	FFBC	FCFD

外部输入 7*7 的 16bit 图片数据

0000	0000	0070	0757	0F1F	1000	0FDF
01A1	0AAA	0BBB	0FCF	0FCF	0FDF	0FCF
0C8C	0FDF	0FCF	0FCF	0D6D	0A8A	0A8A
0E1E	0FDF	0FCF	0FCF	0282	0000	0000
0E2E	1000	0FDF	0A8A	0000	0000	0000

0969	0FDF	0FCF	0A8A	0000	0000	0000
0717	0FDF	0FCF	0A8A	0000	0000	0000

4. 文件说明

名字	功能
TESTBENCH.v	测试
CONV.v	主机

5. 端口说明

Testbench 部分

Signed [6:0] TiData[1:7][1:7]	输入数据 pic
Signed [15:0] ToData[1:5][1:5]	卷积结果
Signed [15:0] TiDataSingle	用于数据传输:测试→DUT
Signed [15:0] ToDataSingle	用于数据传输: DUT →测试
CONV_start	卷积模块开始录入输入数据
CONV_finish	卷积运行结束

CONV 部分

Signed [15:0] CONV_iData	用于数据传输:测试→DUT
Signed [15:0] CONV_oData	用于数据传输: DUT→测试
Signed [15:0] CONV_core[1:9]	卷积核
Signed [15:0] CONV_iArrayData[1:7][1:7]	卷积输入数据
Signed [31:0] CONV_oArrayData[1:5][1:5]	截断前的卷积结果
startCal	卷积模块开始运行
Signed [15:0] CONV_iReCon[1:9]	用于单次卷积的数据阵列
Signed [31:0] CONV_mul[1:9]	单次单点乘积
Signed [31:0] CONV_result	单次卷积结果

6. 工作说明

每次启动后 testbench 读取要卷积的数据,并将此数据传输给 CONV,每次传输一个数据即 16bit。CONV 接收完数据后开始卷积。

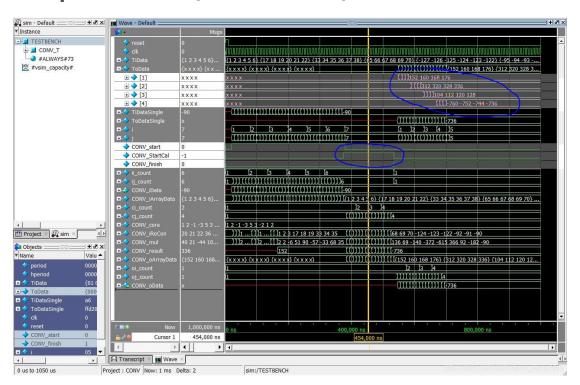
卷积结束后 32bit 截断为 16bit, 把数据传输给 TB, 每次传输一个数据即

16bit。

序号	-	測试流程	控制量
1		TB向 CONV写入数据	TB 置 CONV_start 为 1
2	算	CONV接收完数据后启动卷积运	CONV置 CONV_StartCal为 1
3	据	CONV卷积运算结束开始传输数	CONV 置 CONV_StartCal 为 0 CONV 置 CONV_finish 为 1 TB 置 CONV_start 为 0
4		TB读取 CONV的结果	

7. 仿真结果

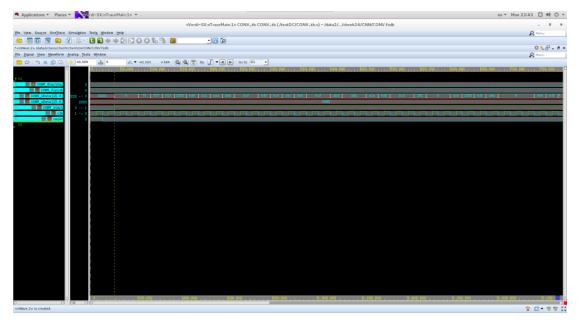
Compile 前仿部分 (波形与代码吻合)



前仿波形: T=20ns, 每 half period (10ns) 时钟翻转一次

T=0 时, reset=1, 持续一周期

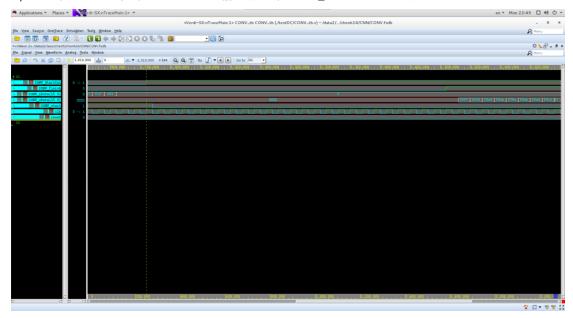
第二周期结束(t=40ns), conv_start=1, 卷积模块开始工作, conv_iData 开始接收 pic 数据



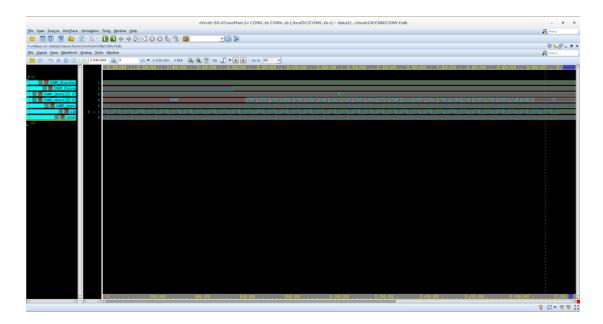
(t=1010ns),conv_startcal=1, 开始正式卷积 MAC 计算, 尽管最后一个 pic 数据还未录入, 但先计算 iArrayData 矩阵的左上角,不影响。

第 51 周期结束(t=1020ns),conv_start=0, pic 所有数据共 49 个录入结束。(t=1510ns), conv_finish=1, 卷积计算结束。

#period, 延迟一个周期 20ns 开始输出数据, conv_oData 开始有数据。



(t=2030ns), 数据输出结束, #period=20ns, clk 停止翻转



DC 仿真 (area report & delay report)

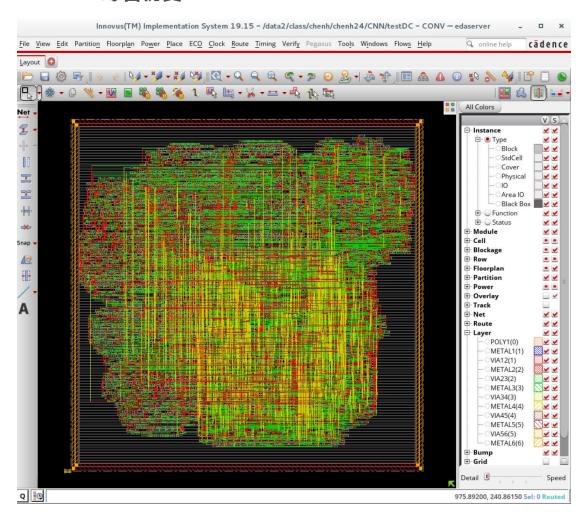
DC area report

```
Report : area
Design : CONV
Version: Q-2019.12-SP5-2
Date : Mon Dec 27 23:10:45 2021
Information: Updating design information... <u>(UID-85)</u>
Warning: Design 'CONV' contains 1 high-fanout nets. A fanout number of 1000 will be used for delay calculati
Library(s) Used:
   slow (File: /data2/class/chenh/chenh45/class_lib/lib/stdcell/aci/sc/synopsys/slow.db)
Number of ports:
                                         2233
Number of nets:
                                        12364
Number of cells:
Number of combinational cells:
Number of sequential cells:
                                         1220
Number of macros/black boxes:
Number of buf/inv:
                                         1470
Number of references:
                              326299.886646
Combinational area:
                               43702.243617
Buf/Inv area:
                               93717.996086
Noncombinational area:
Macro/Black Box area:
                                    0.000000
                            4454449.266724
Net Interconnect area:
Total cell area:
                               420017.882732
Total area:
                              4874467.149456
design_vision>
```

Timing report

Point	Incr	Path	
clock CLK (rise edge) clock network delay (ideal) cj_count_reg[1]/CK (DFFRHQX1) cj_count_reg[1]/CK (DFFRHQX1) U7762/Y (INVX1) U7762/Y (INVX1) U7428/Y (XNOR2X1) U7429/Y (NOR2BX1) U7230/Y (NOR2BX1) U7230/Y (ADI222X1) U7227/Y (NAND3X1) mult_52/B[2] (CONV_DW02_mult_8) mult_52/B[2] (CONV_DW02_mult_8) mult_52/U32/Y (8UFX3) mult_52/S2_4 1/CO (ADDFX2) mult_52/S2_5 1/CO (ADDFX2) mult_52/S2_5 1/CO (ADDFX2) mult_52/S2_7 1/CO (ADDFX2) mult_52/S2_7 1/CO (ADDFX2) mult_52/S2_10_1/CO (ADDFX2) mult_52/S2_10_1/CO (ADDFX2) mult_52/S2_10_1/CO (ADDFX2) mult_52/S2_11_1/CO (ADDFX2) mult_52/S2_11_1/CO (ADDFX2) mult_52/S2_12_1/CO (ADDFX2) mult_52/S2_12_1/CO (ADDFX2) mult_52/S2_13_1/CO (ADDFX2) mult_52/S2_14_1/CO (ADDFX2) mult_52/S2_13_1/CO (ADDFX2) mult_52/S2_14_1/CO (ADDFX2) mult_52/S2_14_1/CO (ADDFX2) mult_52/S2_14_1/CO (ADDFX2) mult_52/S2_14_1/CO (ADDFX2) mult_52/S2_14_1/CO (ADDFX2)	0.00	0.00	
clock CLK (rise edge) clock network delay (ideal)	0.00	0.00	
cj_count_reg[1]/CK (DFFRHQX1)	0.00 #	0.00	r
cj count reg[1]/Q (DFFRHQX1)	0.31	0.31	
U7762/Y (INVX1)	0.26	0.57	
U7428/Y (XN0R2X1)	0.43	1.00	
U7080/Y (INVX1)	0.10	1.10	
U7429/Y (N0R2BX1)	2.42	3.52	
U7230/Y (A0I222X1)	0.35	3.87	f
U7227/Y (NAND3X1)	0.26	4.13	r
mult 52/B[2] (CONV DW02 mult 8)	0.00	4.13	
mult_52/U132/Y (BUFX3)	0.33	4.46	r
mult_52/U3/Y (AND2X2)	0.20	4.66	r
mult_52/S2_4_1/C0 (ADDFX2)	0.57	5.23	r
mult_52/S2_5_1/C0 (ADDFX2)	0.58	5.81	r
mult_52/S2_6_1/C0 (ADDFX2)	0.58	6.38	r
mult_52/S2_7_1/C0 (ADDFX2)	0.58	6.97	
mult_52/U4/Y (AND2X2)	0.17	7.13	
mult_52/S2_9_1/C0 (ADDFX2)	0.57	7.70	
mult_52/S2_10_1/C0 (ADDFX2)	0.58	8.28	
mult_52/S2_11_1/C0 (ADDFX2)	0.58	8.86	
mult_52/S2_12_1/CO (ADDFX2)	0.58	9.44	
mult_52/S2_13_1/CO (ADDFX2)	0.58	10.01	
mace_32, 02_11_1, 00 (NDDI NZ)	0.58	10.59	
mult_52/S4_1/S (ADDFX2)	0.60	11.19	
mult_52/U32/Y (XOR2X1)	0.29	11.48	
mult_52/FS_1/A[14] (CONV_DW01_add_16_DW01_add_18)	0.00 0.15	11.48	
mult_52/FS_1/U99/Y (NAND2X1) mult 52/FS 1/U98/Y (OAI21XL)	0.13	11.63 11.75	
mult 52/FS 1/U14/Y (INVX1)	0.12	11.85	
mult 52/FS 1/SUM[14] (CONV DW01 add 16 DW01 add 18)	0.00	11.85	
mult_52/PRODUCT[16] (CONV_DW01_add_16_bw61_add_16)	0.00	11.85	
add_7_root_add_0_root_add_61_8/B[16] (CONV_DW01_add_7_0			1
ada_,sst_ada_sost_ada_ot_s/b[to] (comv_bwot_add_/_t	0.00	11.85	r
add 7 root add 0 root add 61 8/U1 16/S (ADDFX2)	0.59	12.44	
add_7_root_add_0_root_add_61_8/SUM[16] (CONV_DW01_add_7			•
,	0.00	12.44	f
add 3 root add 0 root add 61 8/B[16] (CONV DW01 add 4 D	W01 add 6		
	0.00	12.44	f
add_3_root_add_0_root_add_61_8/U1_16/S (ADDFX2)	0.59	13.03	f
add_3_root_add_0_root_add_61_8/SUM[16] (CONV_DW01_add_4	LDW01_add	_6)	
	0.00	13.03	f
add_1_root_add_0_root_add_61_8/A[16] (CONV_DW01_add_1_D			
	0.00	13.03	
add_1_root_add_0_root_add_61_8/U1_16/C0 (ADDFX2)	0.50	13.54	
add_1_root_add_0_root_add_61_8/U1_17/C0 (ADDFX2)	0.35	13.89	
add_1_root_add_0_root_add_61_8/U1_18/C0 (ADDFX2)	0.35	14.23	
add_1_root_add_0_root_add_61_8/U1_19/C0 (ADDFX2)	0.35	14.58	
add 7 root add 0 root add 61 8/U1 16/S (ADDFX2)	0.59	12.44	f
add 7 root add 0 root add 61 8/SUM[16] (CONV DW01 add 7			
	0.00	12.44	f
add 3 root add 0 root add 61 8/B[16] (CONV DW01 add 4 DV	w01 add 6)		
	0.00	12.44	f
add_3_root_add_0_root_add_61_8/U1_16/S (ADDFX2)	0.59	13.03	f
add_3_root_add_0_root_add_61_8/SUM[16] (CONV_DW01_add_4	DW01_add	6)	
	0.00	13.03	f
add_1_root_add_0_root_add_61_8/A[16] (CONV_DW01_add_1_DV	w01_add_3)		
	0.00	13.03	f
add_1_root_add_0_root_add_61_8/U1_16/C0 (ADDFX2)	0.50	13.54	f
add_1_root_add_0_root_add_61_8/U1_17/C0 (ADDFX2)	0.35	13.89	
add_1_root_add_0_root_add_61_8/U1_18/C0 (ADDFX2)	0.35	14.23	
add_1_root_add_0_root_add_61_8/U1_19/CO (ADDFX2)	0.35	14.58	
add_1_root_add_0_root_add_61_8/U1_20/C0 (ADDFX2)	0.35	14.93	
add_1_root_add_0_root_add_61_8/U1_21/C0 (ADDFX2)	0.35	15.28	
add_1_root_add_0_root_add_61_8/U1_22/C0 (ADDFX2)	0.35	15.63	
		15.98	
add 1 root add 0 root add 61 8/U1 24/C0 (ADDFX2)		16.33	
add 1 root add 0 root add 61 8/U1 25/C0 (ADDFX2)	0.35	16.68	
add_1_root_add_0_root_add_61_8/U1_26/C0 (ADDFX2)	0.35	17.02	
add 1 root add 0 root add 61 8/U1 27/C0 (ADDFX2)	0.35	17.37	
add_1_root_add_0_root_add_61_8/U1_28/C0 (ADDFX2)	0.35	17.72	
	0.35	18.07	
add_1_root_add_0_root_add_61_8/U1_30/C0 (ADDFX2)	0.36	18.43	
add_1_root_add_0_root_add_61_8/U1_31/Y (XOR3X2)	0.30	18.73	r
add 1 root add 0 root add 61 8/SUM[31] (CONV DW01 add 1	_DW01_add_ 0.00	3) 18.73	r
	0.00		ſ.
add 0 root add 0 root add 61 9/PI313 (CONV DWG) add 6 D	MO1 add 31		
add_0_root_add_0_root_add_61_8/B[31] (CONV_DW01_add_0_D)		18 72	
	0.00	18.73	
add_0_root_add_0_root_add_61_8/U1_31/Y (X0R3X2)	0.00 0.50	19.23	
	0.00 0.50 _DW01_add_	19.23 2)	f
add_0_root_add_0_root_add_61_8/U1_31/Y (XOR3X2) add_0_root_add_0_root_add_61_8/SUM[31] (CONV_DW01_add_0	0.00 0.50	19.23 2)	f
add_0_root_add_0_root_add_61_8/U1_31/Y (X0R3X2)	0.00 0.50 _DW01_add_ 0.00	19.23 2) 19.23	f
add_0_root_add_0_root_add_61_8/U1_31/Y (XOR3X2) add_0_root_add_0_root_add_61_8/SUM[31] (CONV_DW01_add_0_ CONV_oArrayData_reg[1][1][31]/D (EDFFX1) data_arrival_time	0.00 0.50 DW01_add_ 0.00 0.00	19.23 2) 19.23 19.23	f
add_0_root_add_0_root_add_61_8/U1_31/Y (XOR3X2) add_0_root_add_0_root_add_61_8/SUM[31] (CONV_DW01_add_0_ CONV_oArrayData_reg[1][1][31]/D (EDFFX1) data_arrival_time	0.00 0.50 _DW01_add_ 0.00	19.23 2) 19.23 19.23	f
add_0_root_add_0_root_add_61_8/U1_31/Y (XOR3X2) add_0_root_add_0_root_add_61_8/SUM[31] (CONV_DW01_add_0 CONV_oArrayData_reg[1][1][31]/D (EDFFX1) data_arrival_time clock_CLK_(rise_edge) clock_network_delay_(ideal)	0.00 0.50 DW01_add_ 0.00 0.00	19.23 2) 19.23 19.23 19.23 20.00 20.00	f f
add_0_root_add_0_root_add_61_8/U1_31/Y (XOR3X2) add_0_root_add_0_root_add_61_8/SUM[31] (CONV_DW01_add_0] CONV_oArrayData_reg[1][1][31]/D (EDFFX1) data_arrival_time clock CLK (rise_edge) clock_network_delay (ideal) CONV_oArrayData_reg[1][1][31]/CK (EDFFX1)	0.00 0.50 DW01_add_ 0.00 0.00	19.23 2) 19.23 19.23 19.23 20.00 20.00 20.00	f f
add_0_root_add_0_root_add_61_8/U1_31/Y (XOR3X2) add_0_root_add_0_root_add_61_8/SUM[31] (CONV_DW01_add_0_0_0_0_0_0_0_0_0_0_0_0_0_0_0_0_0_0	0.00 0.50 DW01_add_ 0.00 0.00	19.23 2) 19.23 19.23 19.23 20.00 20.00 20.00 19.26	f f
add_0_root_add_0_root_add_61_8/U1_31/Y (XOR3X2) add_0_root_add_0_root_add_61_8/SUM[31] (CONV_DW01_add_0_0_0_0_0_0_0_0_0_0_0_0_0_0_0_0_0_0	0.00 0.50 DW01_add_ 0.00 0.00 20.00 0.00 0.00 -0.74	19.23 2) 19.23 19.23 19.23 20.00 20.00 20.00 19.26 19.26	f f
add_0_root_add_0_root_add_61_8/U1_31/Y (XOR3X2) add_0_root_add_0_root_add_61_8/SUM[31] (CONV_DW01_add_0] CONV_oArrayData_reg[1][1][31]/D (EDFFX1) data arrival time clock CLK (rise edge) clock network delay (ideal) CONV_oArrayData_reg[1][1][31]/CK (EDFFX1) library setup time data required time	0.00 0.50 DW01_add_ 0.00 0.00 20.00 0.00 0.00 -0.74	19.23 2) 19.23 19.23 19.23 20.00 20.00 20.00 19.26 19.26	f f
add_0_root_add_0_root_add_61_8/U1_31/Y (XOR3X2) add_0_root_add_0_root_add_61_8/SUM[31] (CONV_DW01_add_0_0_0_0_0_0_0_0_0_0_0_0_0_0_0_0_0_0	0.00 0.50 0.50 0.00 0.00 0.00 0.00 0.00	19.23 2) 19.23 19.23 19.23 20.00 20.00 20.00 19.26 19.26	f f
add_0_root_add_0_root_add_61_8/U1_31/Y (XOR3X2) add_0_root_add_0_root_add_61_8/SUM[31] (CONV_DW01_add_0] CONV_oArrayData_reg[1][1][31]/D (EDFFX1) data arrival time clock CLK (rise edge) clock network delay (ideal) CONV_oArrayData_reg[1][1][31]/CK (EDFFX1) library setup time data required time	0.00 0.50 0.50 0.00 0.00 0.00 20.00 0.00	19.23 2) 19.23 19.23 19.23 20.00 20.00 20.00 19.26 19.26	f f r

Innovus 综合仿真



optDesign Final Summary Setup views included: setup_view Setup mode | all | reg2reg | default | WNS (ns): | 0.012 | 0.012 | 16.076 | TNS (ns): | 0.000 | 0.000 | 0.000 | Violating Paths: | 0 | 0 | 0 All Paths:| 2472 | 1668 | 2027 Real | Total | Nr nets(terms) | Worst Vio | Nr nets(terms) max_cap 0 (0) 0.000 0 (0) max_tran 0 (0) 0.000 0 (0) max_fanout 0 (0) 0 (0) 0 (0) max_length 0 (0) 0 (0) 0 (0) Density: 50.139% Routing Overflow: 0.00% H and 0.00% V **optDesign ... cpu = 0:00:43, real = 0:00:42, mem = 1218.5M, totSessionCpu=0:14:26 ** *** Finished optDesign *** Info: Destroy the CCOpt slew target map. innovus 9> Start to check current routing status for nets... All nets are already routed correctly. End to check current routing status for nets (mem=1396.0M) _____ timeDesign Summary Hold views included: hold view Hold mode | all | reg2reg | default | WNS (ns): | 0.000 | 0.241 0.000 TNS (ns): | 0.000 | 0.000 0.000 Violating Paths:| 0 | 0 1 0 All Paths:| 2472 | 1668 | 2027 Density: 50.139% Routing Overflow: 0.00% H and 0.00% V Reported timing to dir timingReports

Total CPU time: 2.57 sec Total Real time: 2.0 sec

Total Memory Usage: 1380.410156 Mbytes