US 20190065947A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2019/0065947 A1**

NAKAMURA et al. (43) **Pub. Date: Feb. 28, 2019**

(54) **SEMICONDUCTOR DEVICE AND IMAGE RECOGNITION SYSTEM**

(71) Applicant: **RENESAS ELECTRONICS CORPORATION**, Tokyo (JP)

(72) Inventors: **Atsushi NAKAMURA**, Tokyo (JP); **Akira UTAGAWA**, Tokyo (JP); **Shigeru MATSUO**, Tokyo (JP)

(21) Appl. No.: **16/035,010**

(22) Filed: **Jul. 13, 2018**

(30) **Foreign Application Priority Data**

Aug. 25, 2017 (JP) 2017-162057

Publication Classification

(51) **Int. Cl.**

G06N 3/08	(2006.01)
G06K 9/46	(2006.01)
G06N 3/04	(2006.01)

(52) **U.S. Cl.**

CPC ***G06N 3/08*** (2013.01); ***G06N 3/04*** (2013.01); ***G06K 9/4628*** (2013.01)

(57) **ABSTRACT**

A semiconductor device includes an image recognition device having a convolution arithmetic processing circuit. The convolution arithmetic processing circuit includes a coefficient register where coefficients of an integration coefficient table are set, a product calculation circuit that calculates products of an input image and the coefficients, a channel register where a channel number of the integration coefficient table is set, a channel selection circuit that selects an output destination of a cumulative addition arithmetic operation on the basis of the channel number, and a plurality of output registers that store a result of the cumulative addition arithmetic operation. The integration coefficient table is a table where a plurality of input coefficient tables are integrated and the integration coefficient table has a size of N×N. The product calculation circuit can calculate data of N×N all at once.

IMAGE

0	1	1	0
1	0	1	0
1	0	1	0
0	1	0	0

FILTER

1	0
0	1

FIG. 1

IMAGE

0	1	1	0
1	0	1	0
1	0	1	0
0	1	0	0

FILTER

1	0
0	1

FIG. 2A

0x1	1x0	1	0
1x0	0x1	1	0
1	0	1	0
0	1	0	0

0		

0x1+1x0+1x0+0x1=0

FIG. 2B

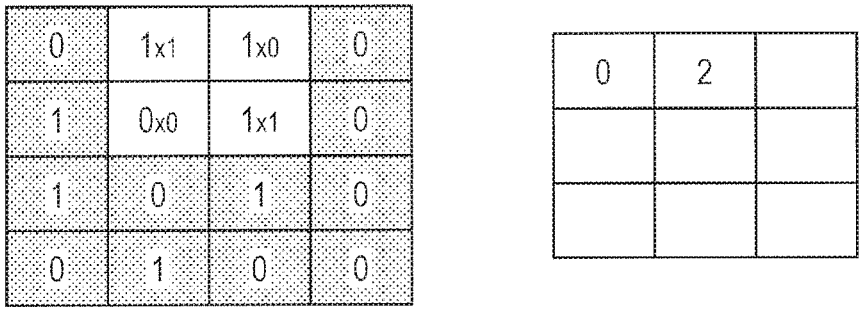

0 1x1 1x0 0
1 0x0 1x1 0
1 0 1 0
0 1 0 0
0 2
1x1+1x0+0x0+1x1=2

FIG. 2C

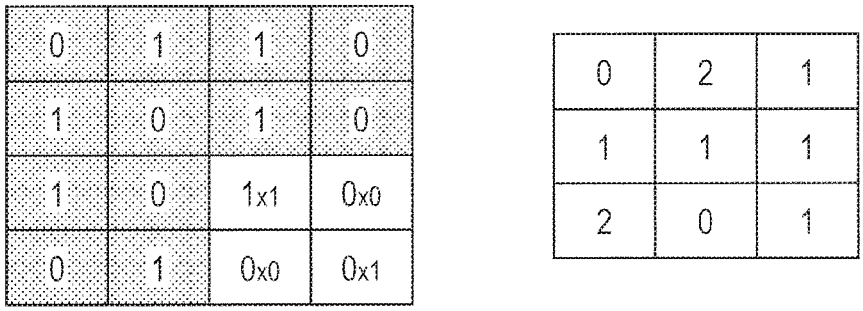

0 1 1 0
1 0 1 0
1 0 1x1 0x0
0 1 0x0 0x1
0 2 1
1 1 1
2 0 1
1x1+0x0+0x0+0x1=1

FIG. 3

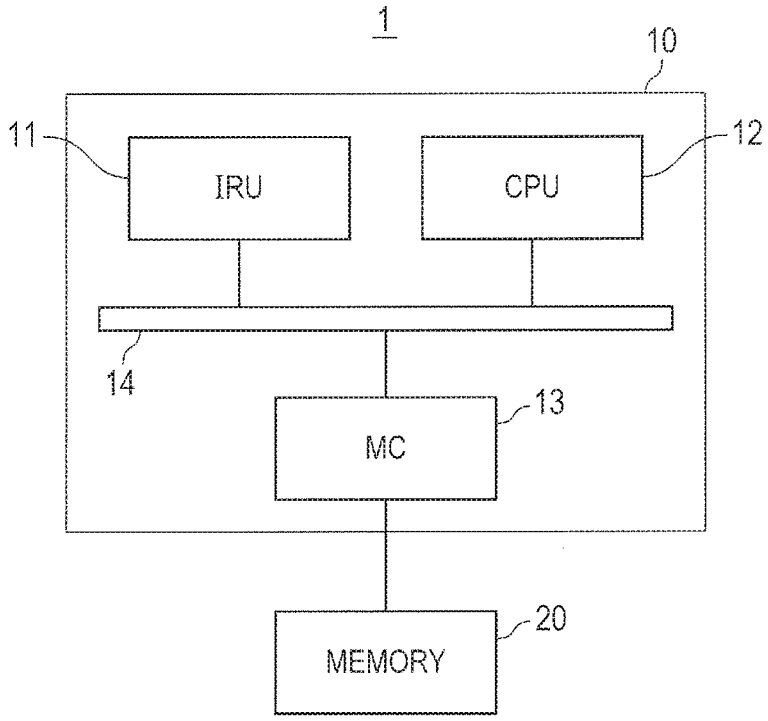


FIG. 4

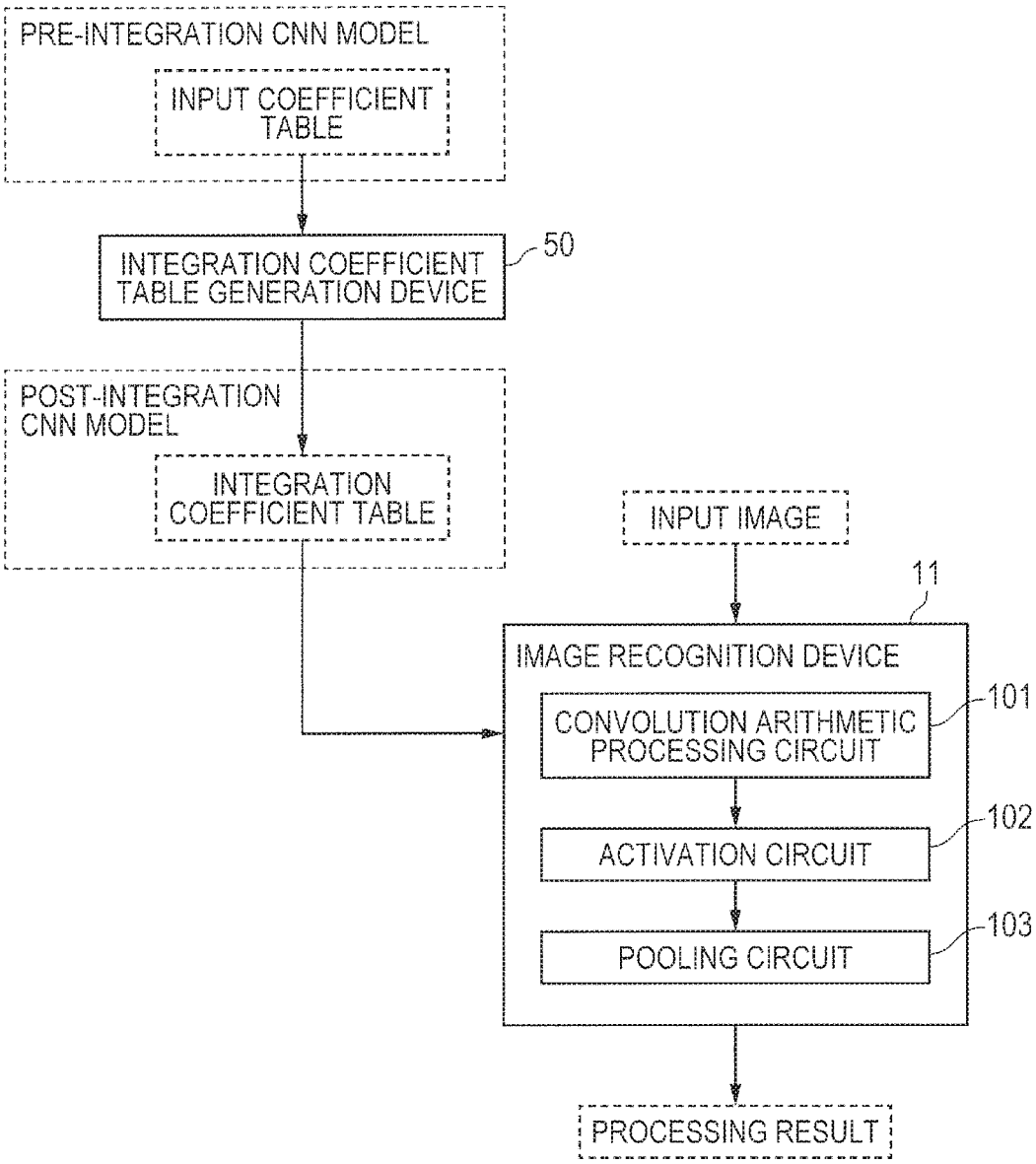

PRE-INTEGRATION CNN MODEL
INPUT COEFFICIENT TABLE
INTEGRATION COEFFICIENT TABLE GENERATION DEVICE
50
POST-INTEGRATION CNN MODEL
INTEGRATION COEFFICIENT TABLE
INPUT IMAGE
11
IMAGE RECOGNITION DEVICE
CONVOLUTION ARITHMETIC PROCESSING CIRCUIT
101
ACTIVATION CIRCUIT
102
POOLING CIRCUIT
103
PROCESSING RESULT

FIG. 5

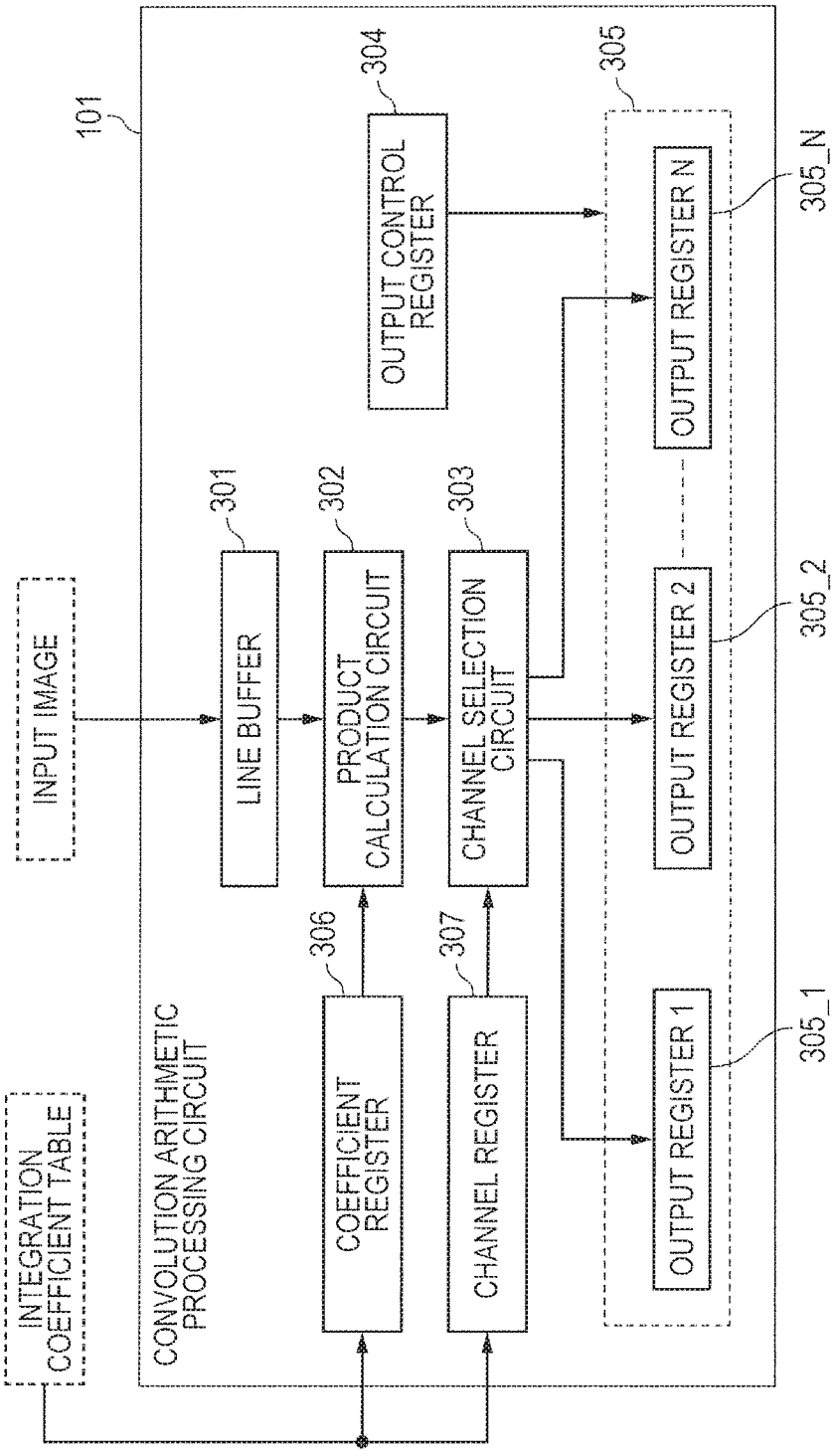

101
INPUT IMAGE
INTEGRATION COEFFICIENT TABLE
CONVOLUTION ARITHMETIC PROCESSING CIRCUIT
LINE BUFFER
301
PRODUCT CALCULATION CIRCUIT
302
CHANNEL SELECTION CIRCUIT
303
OUTPUT CONTROL REGISTER
304
305
OUTPUT REGISTER 1
OUTPUT REGISTER 2
OUTPUT REGISTER N
305_1
305_2
305_N
COEFFICIENT REGISTER
306
CHANNEL REGISTER
307

FIG. 6

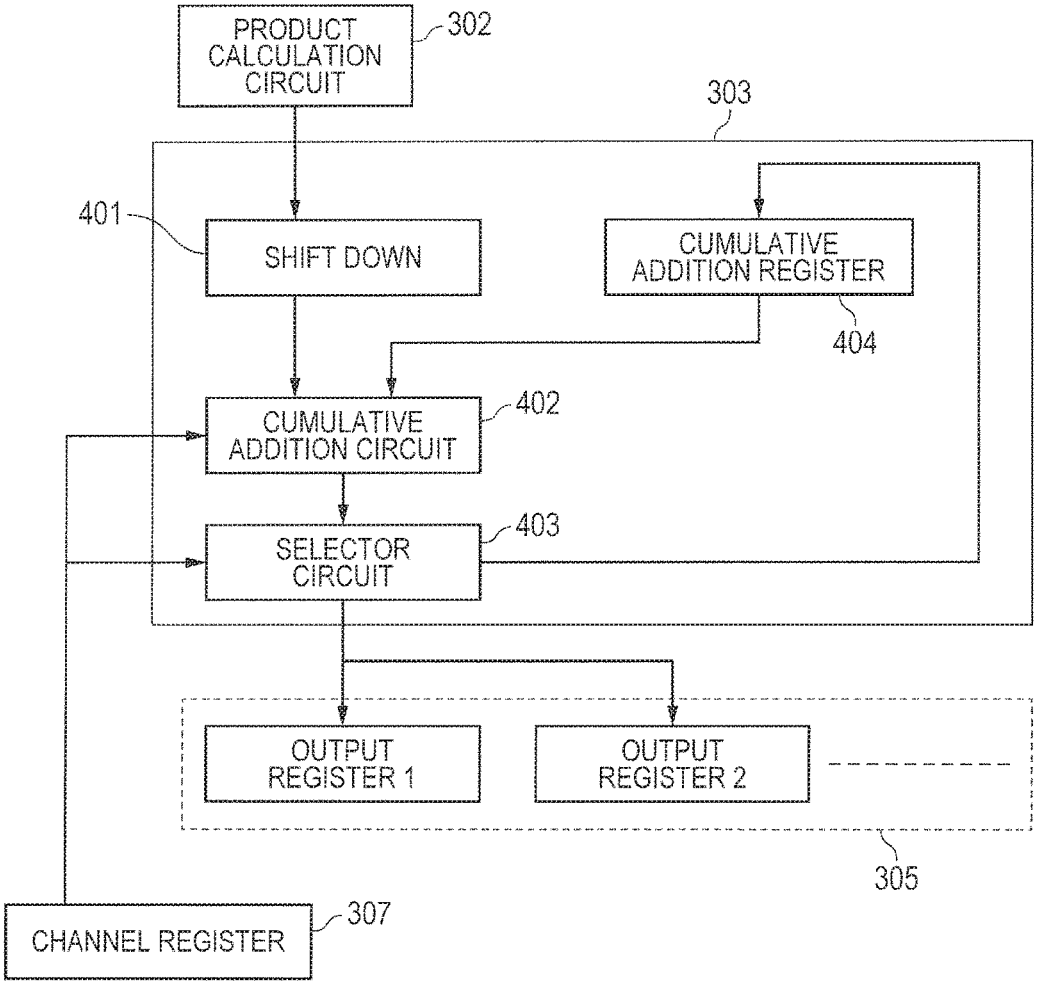

PRODUCT CALCULATION CIRCUIT
302
303
401
SHIFT DOWN
CUMULATIVE ADDITION REGISTER
404
CUMULATIVE ADDITION CIRCUIT
402
SELECTOR CIRCUIT
403
OUTPUT REGISTER 1
OUTPUT REGISTER 2
305
CHANNEL REGISTER
307

FIG. 7

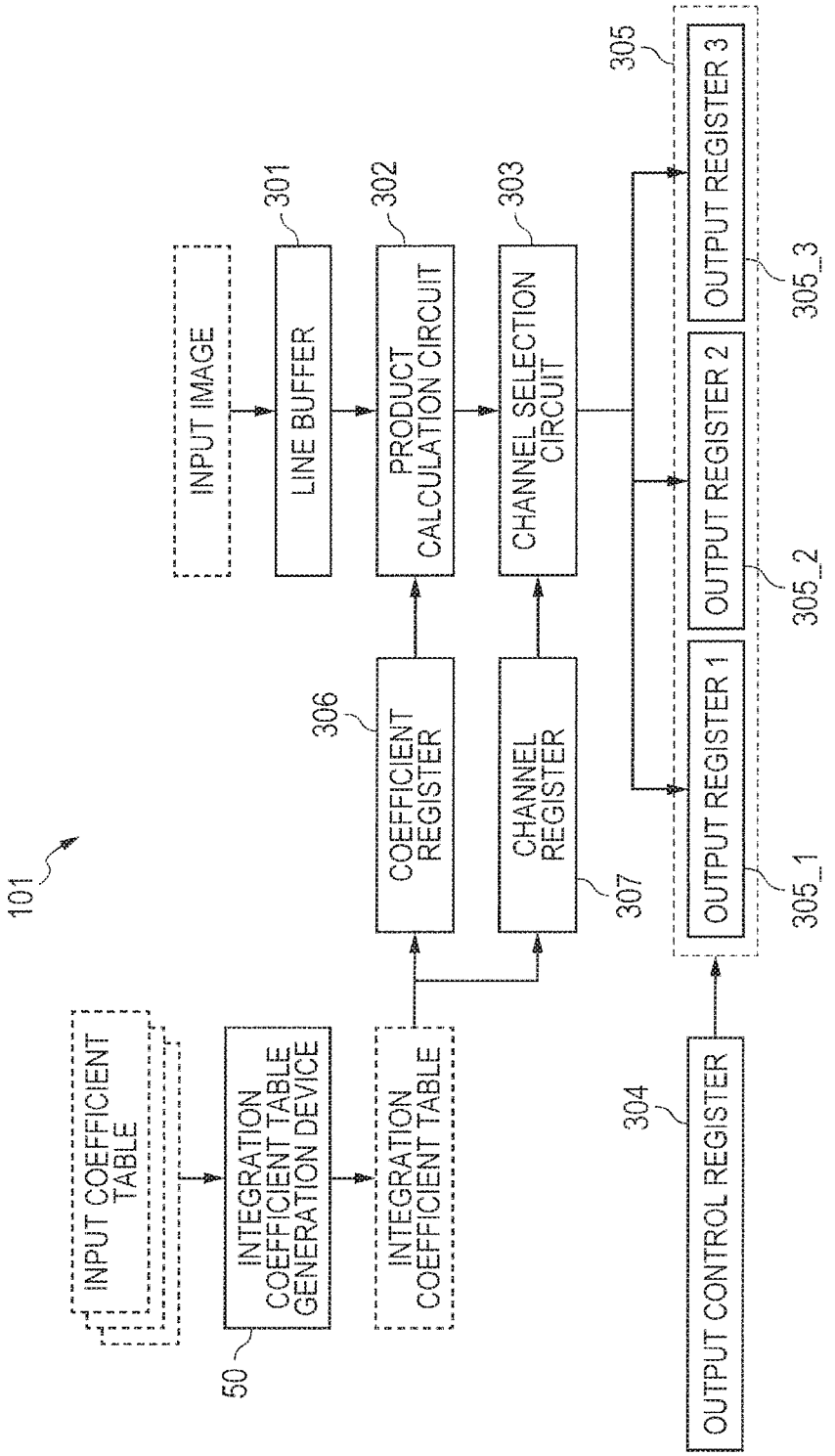

101
INPUT COEFFICIENT TABLE
INTEGRATION COEFFICIENT TABLE GENERATION DEVICE
50
INTEGRATION COEFFICIENT TABLE
INPUT IMAGE
LINE BUFFER
301
PRODUCT CALCULATION CIRCUIT
302
CHANNEL SELECTION CIRCUIT
303
COEFFICIENT REGISTER
306
CHANNEL REGISTER
307
OUTPUT CONTROL REGISTER
304
305
OUTPUT REGISTER 1
305_1
OUTPUT REGISTER 2
305_2
OUTPUT REGISTER 3
305_3

FIG. 8

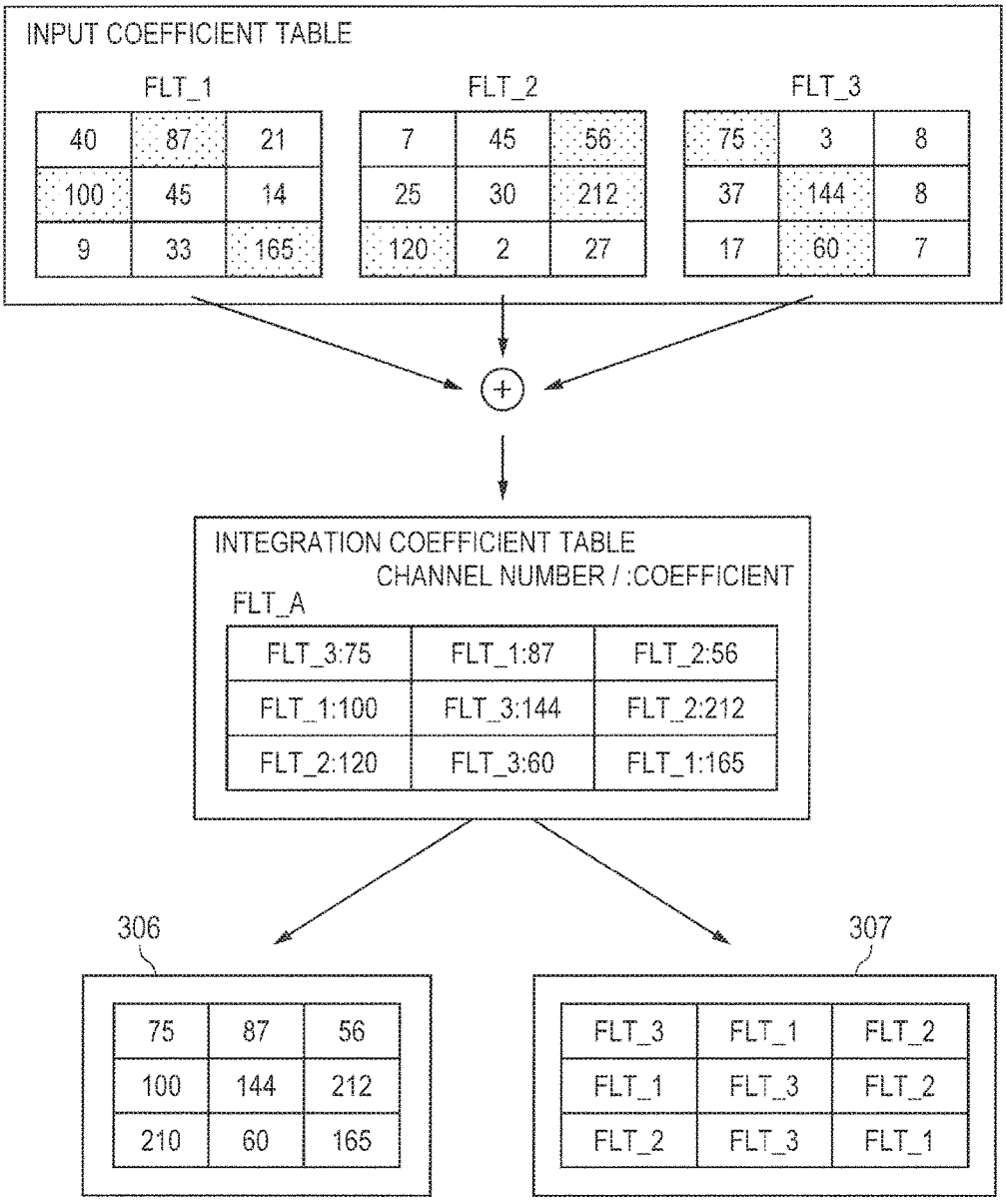

INPUT COEFFICIENT TABLE
FLT_1
40 87 21
100 45 14
9 33 165
FLT_2
7 45 56
25 30 212
120 2 27
FLT_3
75 3 8
37 144 8
17 60 7
+
INTEGRATION COEFFICIENT TABLE
CHANNEL NUMBER / :COEFFICIENT
FLT_A
FLT_3:75 FLT_1:87 FLT_2:56
FLT_1:100 FLT_3:144 FLT_2:212
FLT_2:120 FLT_3:60 FLT_1:165
306
75 87 56
100 144 212
210 60 165
307
FLT_3 FLT_1 FLT_2
FLT_1 FLT_3 FLT_2
FLT_2 FLT_3 FLT_1

FIG. 9

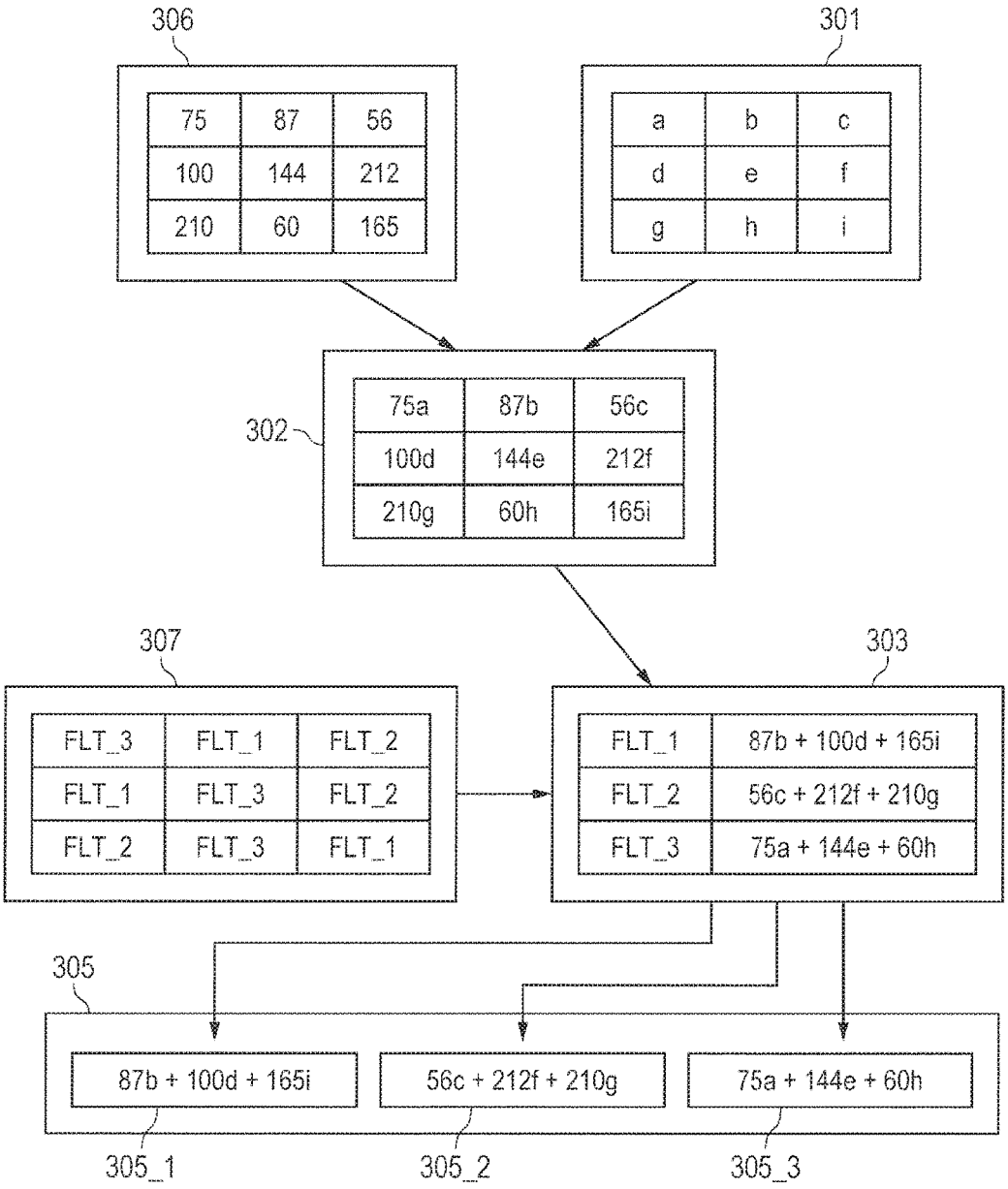

306
75 87 56
100 144 212
210 60 165
301
a b c
d e f
g h i
302
75a 87b 56c
100d 144e 212f
210g 60h 165i
307
FLT_3 FLT_1 FLT_2
FLT_1 FLT_3 FLT_2
FLT_2 FLT_3 FLT_1
303
FLT_1 87b + 100d + 165i
FLT_2 56c + 212f + 210g
FLT_3 75a + 144e + 60h
305
87b + 100d + 165i
56c + 212f + 210g
75a + 144e + 60h
305_1
305_2
305_3

FIG. 10

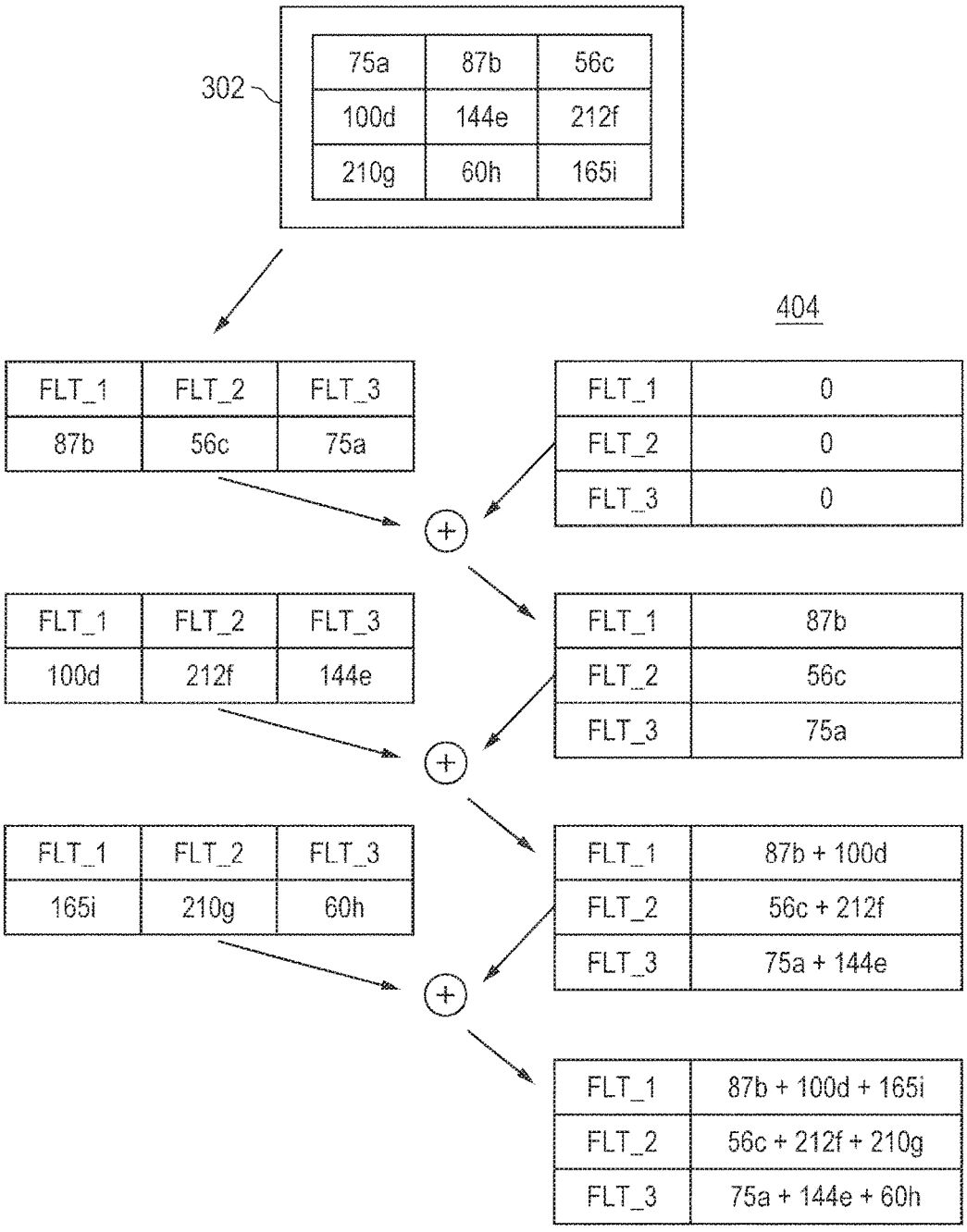

302
75a
87b
56c
100d
144e
212f
210g
60h
165i
404
FLT_1
FLT_2
FLT_3
87b
56c
75a
0
0
0
100d
212f
144e
87b
56c
75a
165i
210g
60h
87b + 100d
56c + 212f
75a + 144e
87b + 100d + 165i
56c + 212f + 210g
75a + 144e + 60h

FIG. 11

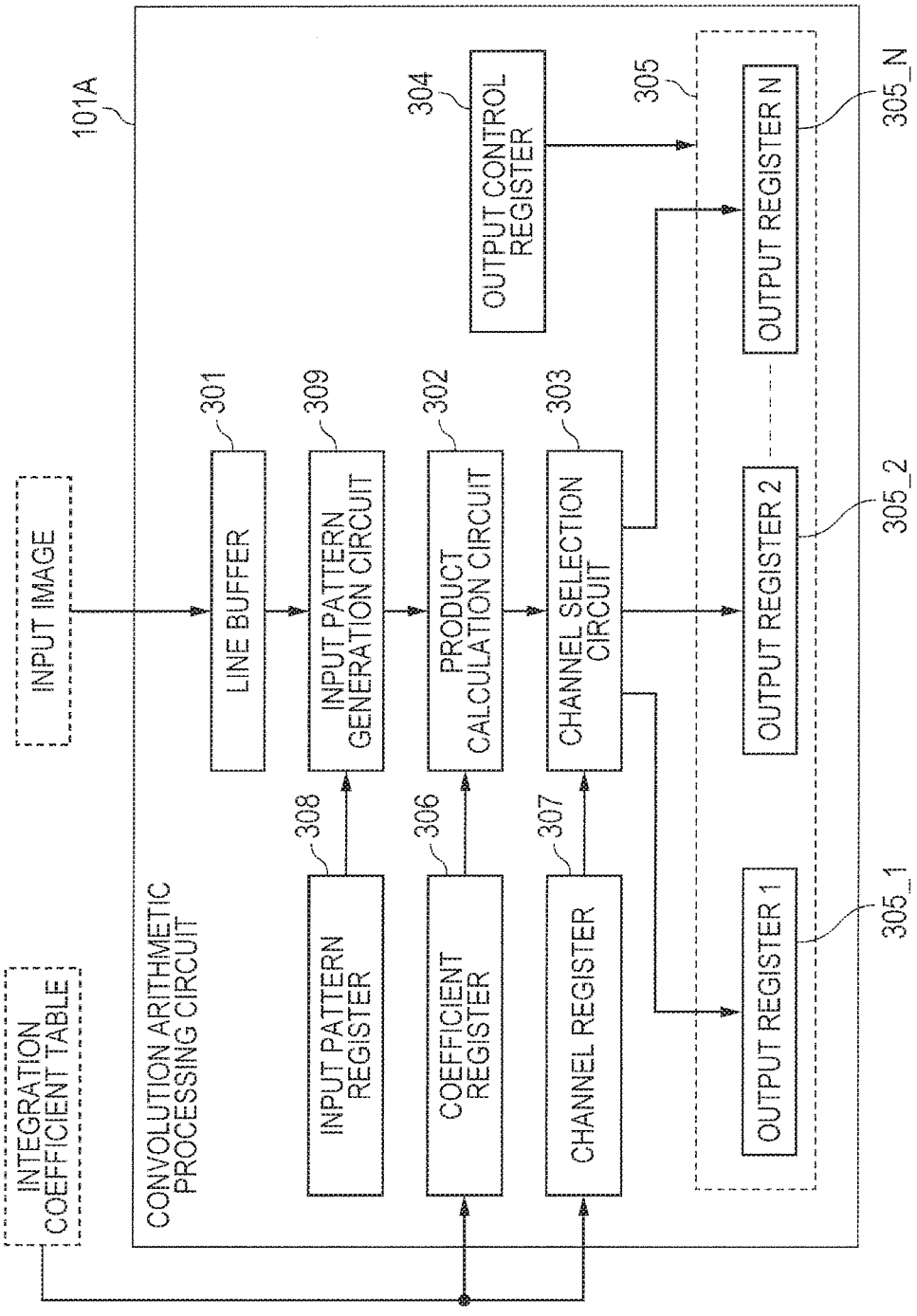

101A
CONVOLUTION ARITHMETIC PROCESSING CIRCUIT
INPUT IMAGE
INTEGRATION COEFFICIENT TABLE
LINE BUFFER
301
INPUT PATTERN GENERATION CIRCUIT
309
PRODUCT CALCULATION CIRCUIT
302
CHANNEL SELECTION CIRCUIT
303
OUTPUT CONTROL REGISTER
304
305
OUTPUT REGISTER N
305_N
OUTPUT REGISTER 2
305_2
OUTPUT REGISTER 1
305_1
INPUT PATTERN REGISTER
308
COEFFICIENT REGISTER
306
CHANNEL REGISTER
307

FIG. 12

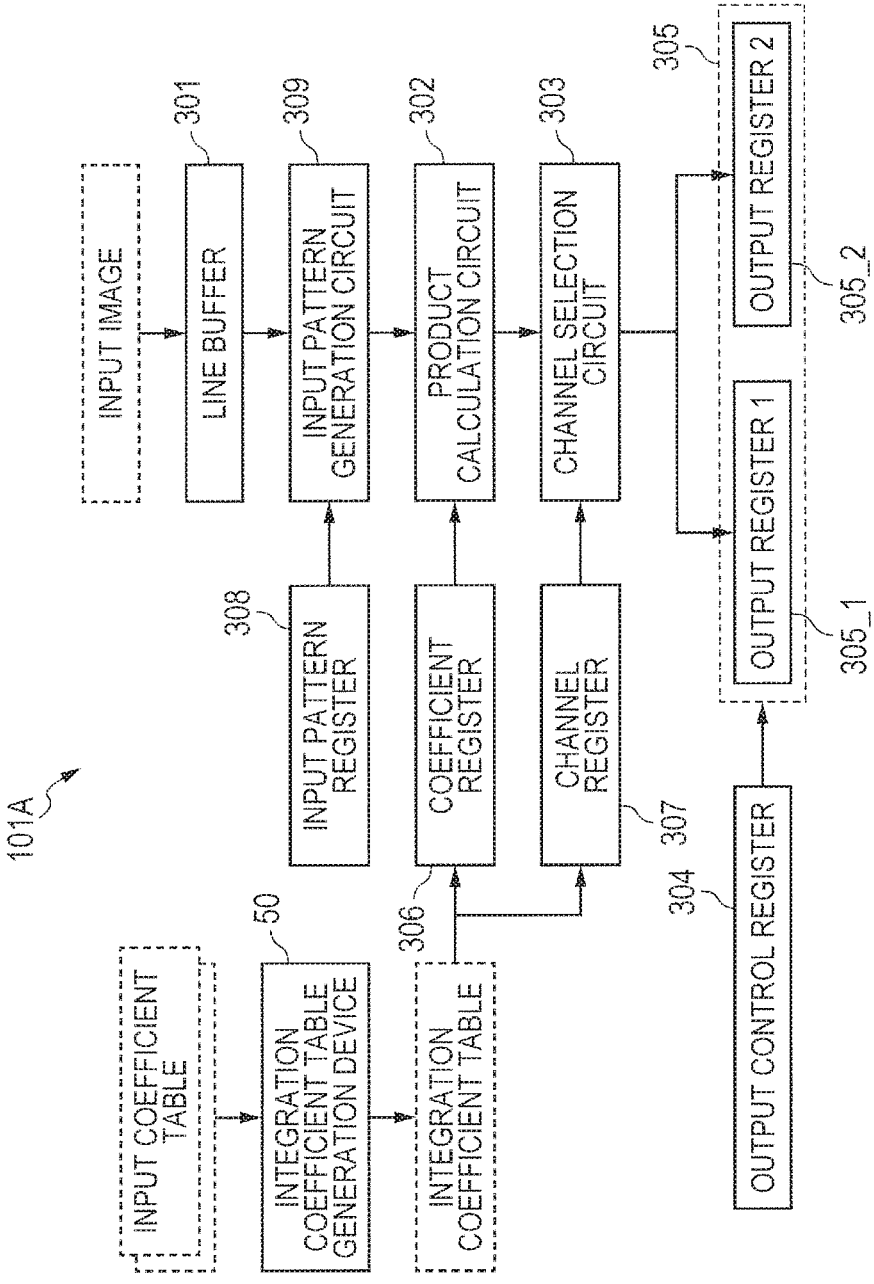

101A
INPUT IMAGE
301
LINE BUFFER
309
INPUT PATTERN GENERATION CIRCUIT
302
PRODUCT CALCULATION CIRCUIT
303
CHANNEL SELECTION CIRCUIT
305
OUTPUT REGISTER 2
305_2
OUTPUT REGISTER 1
305_1
308
INPUT PATTERN REGISTER
COEFFICIENT REGISTER
CHANNEL REGISTER
307
304
OUTPUT CONTROL REGISTER
306
INTEGRATION COEFFICIENT TABLE
50
INTEGRATION COEFFICIENT TABLE GENERATION DEVICE
INPUT COEFFICIENT TABLE

FIG. 13

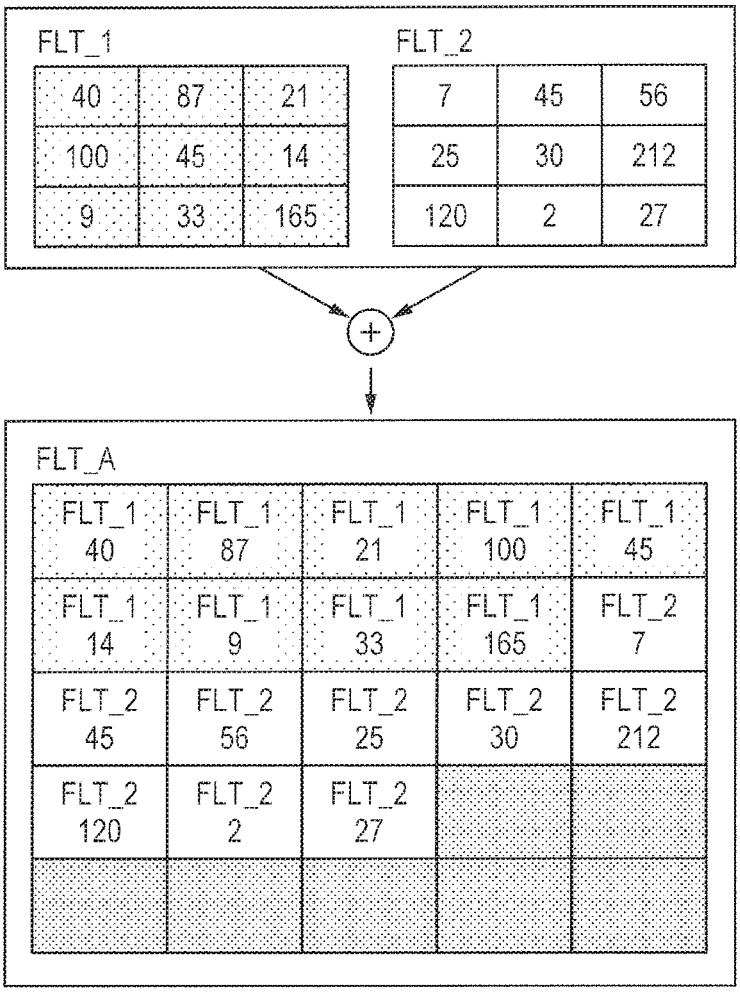
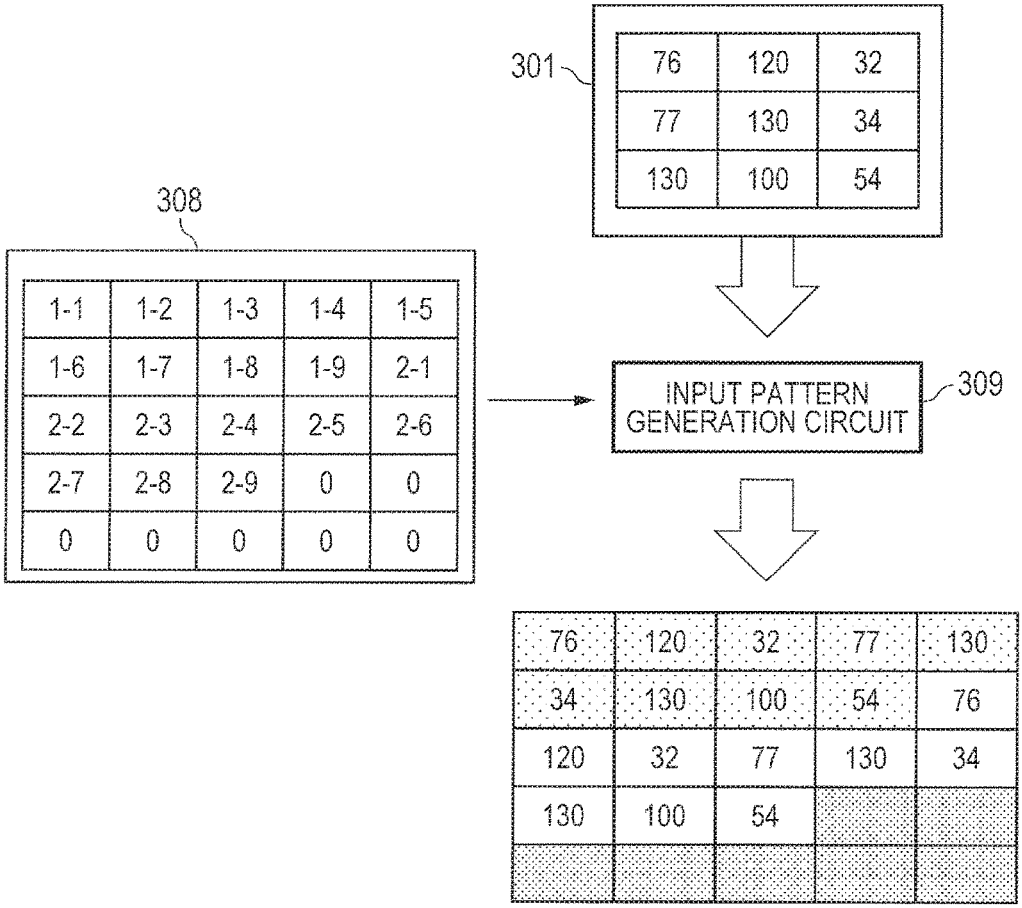

FLT_1
40 87 21
100 45 14
9 33 165
FLT_2
7 45 56
25 30 212
120 2 27
+
FLT_A
FLT_1 40
FLT_1 87
FLT_1 21
FLT_1 100
FLT_1 45
FLT_1 14
FLT_1 9
FLT_1 33
FLT_1 165
FLT_2 7
FLT_2 45
FLT_2 56
FLT_2 25
FLT_2 30
FLT_2 212
FLT_2 120
FLT_2 2
FLT_2 27

FIG. 14


301
76 120 32
77 130 34
130 100 54
308
1-1 1-2 1-3 1-4 1-5
1-6 1-7 1-8 1-9 2-1
2-2 2-3 2-4 2-5 2-6
2-7 2-8 2-9 0 0
0 0 0 0 0
INPUT PATTERN GENERATION CIRCUIT
309
76 120 32 77 130
34 130 100 54 76
120 32 77 130 34
130 100 54

SEMICONDUCTOR DEVICE AND IMAGE RECOGNITION SYSTEM

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] The disclosure of Japanese Patent Application No. 2017-162057 filed on Aug. 25, 2017 including the specification, drawings and abstract is incorporated herein by reference in its entirety.

BACKGROUND

[0002] The present disclosure relates to a semiconductor device and, for example, can be applied to a semiconductor device including an image recognition device.

[0003] Since Convolutional Neural Network (CNN) won first place in an international image classification contest (ImageNet Large Scale Visual Recognition Competition: ILSVRC) in 2012, the CNN has been used in many image recognition processes and its usefulness has been recognized. In recent years, there is a growing demand to use the CNN not only in a workstation with rich resources but also in a built-in type device. The CNN is disclosed in, for example, Japanese Unexamined Patent Application Publication No. 2008-310700 and Japanese Unexamined Patent Application Publication No. 2016-99707.

SUMMARY

[0004] However, the CNN is required to repeat a convolution arithmetic operation of an input image and coefficients the number of times corresponding to the number of channels, so that an operation amount is extremely large. Therefore, it is difficult to implement the CNN on a built-in type SoC (Silicon on Chip) whose resource is limited. The other problems and novel features will become apparent from the description of the present specification and the accompanying drawings.

[0005] A brief outline of a typical invention of the present disclosure is as follows. A semiconductor device has a function to perform cumulative addition for each channel and performs a convolution arithmetic operation in parallel in a plurality of channels.

[0006] According to the semiconductor device described above, it is possible to reduce the operation amount of the convolution arithmetic operation.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] FIG. **1** is a diagram for explaining a concept of a convolution arithmetic operation.

[0008] FIG. **2**A is a diagram for explaining a concept of the convolution arithmetic operation.

[0009] FIG. **2**B is a diagram for explaining a concept of the convolution arithmetic operation.

[0010] FIG. **2**C is a diagram for explaining a concept of the convolution arithmetic operation.

[0011] FIG. **3** is a diagram showing a configuration example of an image recognition system.

[0012] FIG. **4** is a diagram showing a flow of data of the image recognition system in FIG. **3**.

[0013] FIG. **5** is a diagram for explaining a configuration example of a convolution arithmetic processing circuit in FIG. **4**.

[0014] FIG. **6** is a diagram for explaining a configuration example of a channel selection circuit in FIG. **5**.

[0015] FIG. **7** is a diagram for explaining a configuration of a 3×3 convolution arithmetic processing circuit.

[0016] FIG. **8** is a schematic diagram for explaining an operation of an integration coefficient table generation device in FIG. **7**.

[0017] FIG. **9** is a conceptual diagram for explaining an operation of the convolution arithmetic processing circuit in FIG. **7**.

[0018] FIG. **10** is a conceptual diagram for explaining an operation of a cumulative addition circuit in FIG. **7**.

[0019] FIG. **11** is a diagram for explaining another configuration example of the convolution arithmetic processing circuit in FIG. **4**.

[0020] FIG. **12** is a diagram for explaining a configuration of a 5×5 convolution arithmetic processing circuit.

[0021] FIG. **13** is a schematic diagram for explaining an operation of an integration coefficient table generation device in FIG. **12**.

[0022] FIG. **14** is a schematic diagram for explaining an operation of an input pattern generation circuit in FIG. **12**.

DETAILED DESCRIPTION

[0023] First, a convolution arithmetic operation of CNN will be described with reference to FIGS. **1** and **2**A to **2**C. FIG. **1** is a diagram for explaining a concept of the convolution arithmetic operation. FIGS. **2**A to **2**C are diagrams for explaining a concept of the convolution arithmetic operation and are diagrams showing a product sum arithmetic operation of pixel values of an image and coefficients of a filter.

[0024] For ease of description, when an image is simplified and represented by numbers, the image can be considered as a matrix of height×width. As shown in the left of FIG. **1**, an input image is represented by, for example, 4×4 binary values. A 2×2 filter as shown in the right of FIG. **1** is applied to such a matrix and a convolution arithmetic operation is performed.

[0025] As shown in FIG. **2**A, the filer is overlapped on a most upper left portion of the image, and values in the same positions are multiplied together. The results of the multiplication are summed up and outputted. As shown in FIG. **2**B, the filter is shifted horizontally by one pixel, and the values in the same positions are multiplied together in the same manner. The results of the multiplication are summed up and outputted. Thereafter, as shown in FIG. **2**C, the values of the filter and the input image are multiplied together and summed up until the filter reaches the lower right.

[0026] When the convolution arithmetic operation is performed, it is required that the convolution arithmetic operation of the input image and the coefficients of the filter is repeated for the number of times corresponding to the number of channels (the number of filters).

[0027] An embodiment has a function to perform cumulative addition for each channel and performs a convolution arithmetic operation in parallel in a plurality of channels. It is possible to simultaneously and collectively perform the convolution arithmetic operation on channels, the number of which is the same as the number of blocks where the cumulative addition can be performed for each channel, so that processing time can be reduced.

[0028] It is possible to perform the convolution arithmetic operation of a plurality of channels at one time by combining the convolution arithmetic processing circuit and an integration coefficient table. It is possible to realize the CNN by

a small number of product sum arithmetic operations without performing useless product sum arithmetic operations. By using an input pattern generation circuit, it is possible to perform an arbitrary convolution arithmetic operation without depending on a filer size.

[0029] Hereinafter, an embodiment and a modified example will be described with reference to the drawings. However, in the description below, the same components are denoted by the same reference numerals and the description thereof may be omitted.

Embodiment

[0030] FIG. **3** is a block diagram showing a configuration example of an image recognition system. FIG. **4** is a diagram showing a flow of data of the image recognition system in FIG. **3**.

[0031] As shown in FIG. **3**, an image processing system **1** includes a semiconductor device **10** and a semiconductor storage device (MEMORY) **20**. The semiconductor device **10** includes an image recognition device (IRU) **11**, a central processing unit (CPU) **12**, a control circuit (MC) **13** of the semiconductor storage device **20**, and a bus that couples these, in a single semiconductor chip. The semiconductor storage device **20** is, for example, a DDR-SDRAM.

[0032] As shown in FIG. **4**, an integration coefficient table generation device **50** integrates an input coefficient table included in a pre-integration CNN model stored in a storage device, generates an integration coefficient table of a post-integration CNN model, and stores the integration coefficient table in the semiconductor storage device **20**. The image recognition device **11** of the image processing system **1** processes an input image stored in the semiconductor storage device **20** on the basis of the integration coefficient table and stores a processing result in the semiconductor storage device **20**.

[0033] The input image is an image acquired by an image pickup device (not shown in the drawings) having an optical system (lens) and an image pickup element. The input image is all or part of the image acquired by the image pickup device. The input image is stored in the semiconductor storage device **20**.

[0034] The image recognition device **11** performs calculation by using the integration coefficient table of the post-integration CNN model and the input image and outputs a processing result. The image recognition device **11** is a device that performs an arithmetic operation required in the post-integration CNN model.

[0035] The processing result is an output of the image recognition device **11**. For example, the processing result is a classification result or a segmentation result of the input image or a detection result of a specific object. The processing result is stored in the semiconductor storage device **20**.

[0036] The pre-integration CNN model indicates a general CNN model. For example, the pre-integration CNN model is AlexNet, VGGNet, SegNet, or R-CNN. The input coefficient table includes coefficients used for the convolution arithmetic operation among coefficients included in the pre-integration CNN model.

[0037] The integration coefficient table generation device **50** is a device that analyzes the input coefficient table and generates an integration coefficient table where processing can be performed by using a smaller number of convolution arithmetic operations, and is configured by, for example, a central processing unit and a storage device that stores a program of the central processing unit and the like.

[0038] The post-integration CNN model is a model where the number of convolution arithmetic operations of the pre-integration CNN model is reduced. The integration coefficient table is an output of the integration coefficient table generation device **50**.

[0039] The image recognition device **11** includes a convolution arithmetic processing circuit **101**, an activation circuit **102**, and a pooling circuit **103**, and can cumulatively add product calculation results for each channel. The convolution arithmetic processing circuit **101** adds results of convolution arithmetic operations for each channel number corresponding to each coefficient. The activation circuit **102** and the pooling circuit **103** are circuits that perform activation processing and pooling processing, respectively. For the activation processing, for example, a logistic sigmoid function and a ReLU function (Rectified Linear Units), which are well known, are used. The pooling processing is a method that reduces a size of a large image while keeping important information. In the pooling processing, for example, a well-known maximum pooling function that segments an image into small windows and outputs a maximum value of data inputted from each segmented window and a well-known average pooling function that outputs an average value of inputted data are used.

[0040] FIG. **5** is a block diagram for explaining a configuration example of the convolution arithmetic processing circuit. FIG. **6** is a block diagram for explaining a configuration example of a channel selection circuit in FIG. **5**.

[0041] The integration coefficient table is a table in which coefficients used in the convolution arithmetic operation are stored. A channel number is added to a coefficient. The integration coefficient table generation device **50** generates the integration coefficient table from a plurality of input coefficient tables.

[0042] The convolution arithmetic processing circuit **101** outputs results of a plurality of number of times of convolution arithmetic operations from one convolution arithmetic operation. The convolution arithmetic processing circuit **101** includes a line buffer **301**, a product calculation circuit **302**, a channel selection circuit **303**, an output control register **304**, output registers **305** (a first output register **305_1**, a second output register **305_2**, . . . , and an Nth output register **305**_N), a coefficient register **306**, and a channel register **307**. The line buffer **301** acquires input image data for a specified number of lines from the semiconductor storage device **20**.

[0043] The product calculation circuit **302** calculates a product of the coefficient register **306** and an input image from the line buffer **301**. The channel selection circuit **303** adds a channel number registered in the channel register **307** to a product calculation result, cumulatively adds the product calculation result, and outputs the product calculation result to an appropriate output register **305** according to the channel number in accordance with the output control register **304**. The output control register **304** associates the channel number with an output register number.

[0044] As shown in FIG. **6**, the channel selection circuit **303** includes a shift down circuit **401**, a cumulative addition circuit **402**, a selector circuit **403**, and a cumulative addition register **404**. The shift down circuit **401** shifts down the product calculation result by a specified number of bits and then outputs the product calculation result. The cumulative

addition circuit **402** adds results outputted by the cumulative addition register **404** and the shift down circuit **401** for each channel and outputs the added results. The selector circuit **403** outputs the output of the cumulative addition circuit **402** to a specified output register **305**.

[0045] The output control register **304**, the coefficient register **306**, and the channel register **307** are set by the CPU **12**.

[0046] An example where the integration coefficient table generation device **50** generates one type of integration coefficient table from three types of input coefficient tables and performs the convolution arithmetic operation by using the integration coefficient table will be described with reference to FIGS. **7** to **10**. FIG. **7** is a block diagram for explaining a configuration of a 3×3 convolution arithmetic processing circuit. FIG. **8** is a schematic diagram for explaining an operation of the integration coefficient table generation device in FIG. **7**. FIG. **8** represents a case where the integration coefficient table generation device generates one integration coefficient table from the input coefficient tables for three channels. FIG. **9** is a conceptual diagram for explaining an operation of the convolution arithmetic processing circuit in FIG. **7**. FIG. **10** is a conceptual diagram for explaining an operation of the cumulative addition circuit in FIG. **7**.

[0047] The convolution arithmetic processing circuit in FIG. **7** performs one 3×3 convolution arithmetic operation for three channels by using a 3×3 convolution arithmetic operation circuit.

[0048] Here, a coefficient of a coefficient table used in the convolution is defined as $w_{i,j,chanel}$, the table is defined as $C_{l,w,channel}$, and the following equation is established.

[Formula 1]

$$C_{l,w,channel}=\begin{bmatrix} w_{0,0,channel} & w_{0,1,channel} & \cdots & \cdots \\ w_{1,0,channel} & w_{1,1,channel} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \\ \vdots & \vdots & & w_{l-1,w-1,channel} \end{bmatrix} \quad [数\ 1]$$

[0049] Here, i,j represent integers, channel represents a channel number, l represents the width of the table, w represents the height of the table. A coefficient of the integration coefficient table is represented as follows:

[数 2]

$$w_{i,jFLT_A}=\mathrm{MAX}(w_{i,j,FLT_1},w_{i,j,FLT_2}mw_{i,j,FLT_3}) \quad [\text{Formula 2}]$$

[0050] Here, MAX(. . .) is a function that returns the maximum value of arguments.

[0051] Further, l_{i,j,FLT_A} is defined as a character string that represents a type of a table used as the integration coefficient table, and the following equation is established.

[数 3]

$$l_{i,j,FLT_A}=\mathrm{CHANNEL}(w_{i,j,FLT_1},w_{i,j,FLT_2},w_{i,j,FLT_3}) \quad [\text{Formula 3}]$$

[0052] Here, CHANNEL (. . .) is a function representing a channel number that returns the maximum value of arguments. At this time, $C_{3,3,FL_A}$ of the integration coefficient table is represented as follows:

[Formula 4]

$$C_{3,3,FLT_A}=\begin{bmatrix} l_{0,0,FLT_A}: & l_{0,1,FLT_A}: & l_{0,2,FLT_A}: \\ w_{0,0,FLT_A} & w_{0,1,FLT_A} & w_{0,2,FLT_A} \\ l_{1,0,FLT_A}: & l_{1,1,FLT_A}: & l_{1,2,FLT_A}: \\ w_{0,0,FLT_A} & w_{1,1,FLT_A} & w_{1,2,FLT_A} \\ l_{2,0,FLT_A}: & l_{2,1,FLT_A}: & l_{2,2,FLT_A}: \\ w_{2,0,FLT_A} & w_{2,1,FLT_A} & w_{2,2,FLT_A} \end{bmatrix} \quad [数\ 4]$$

[0053] FIG. **8** schematically represents how the integration coefficient table is generated when specifically l=3 and w=3. The integration coefficient table generation device **50** processes three types of input coefficient tables (FLT_**1**, FLT_**2**, and FLT_**3**) and outputs an integration coefficient table (FLT_A). Largest coefficients of those of the input coefficient tables (FLT_**1**, FLT_**2**, and FLT_**3**) are outputted to the integration coefficient table. Each of the largest coefficients is shaded. The maximum value of the upper left coefficients is "75" of FLT_**3**, the maximum value of the upper center coefficients is "87" of FLT_**1**, and the maximum value of the upper right coefficients is "56" of FLT_**2**. The maximum value of the middle left coefficients is "100" of FLT_**1**, the maximum value of the middle center coefficients is "144" of FLT_**3**, and the maximum value of the middle right coefficients is "212" of FLT_**2**. The maximum value of the lower left coefficients is "120" of FLT_**2**, the maximum value of the lower center coefficients is "60" of FLT_**3**, and the maximum value of the lower right coefficients is "165" of FLT_**1**. In the integration coefficient table (FLT_A) where the above maximum values are integrated, the shaded portions of each of the input coefficient tables (FLT_**1**, FLT_**2**, and FLT_**3**) are arranged. In addition to the coefficients, the channel numbers are also stored in the integration coefficient table (FLT_A).

[0054] As shown in FIG. **8**, the convolution arithmetic processing circuit **101** reads the coefficients and the channel numbers from the integration coefficient table and sets their values in the coefficient register **306** and the channel register **307**, respectively.

[0055] As shown in FIG. **9**, the product calculation circuit **302** calculates products of the values set in the coefficient register **306** and the values set in the line buffer **301** all at once by a 3×3 product calculation circuit. The channel selection circuit **303** sets outputs of the cumulative addition circuit **402** to the output registers **305** on the basis of the values (channel numbers) set in the channel register **307**. The output control register **304** sets a correspondence between the output registers **305** and the channel register **307**. In this example, the first output register **305_1** corresponds to FLT_**1**, the second output register **305_2** corresponds to FLT_**2**, and the third output register **305_3** corresponds to FLT_**3**. The convolution arithmetic processing circuit **101** outputs results of performing a convolution arithmetic operation by using the coefficients of FLT_**1**, FLT_**2**, and FLT_**3** to the first output register **305_1**, the second output register **305_2**, and the third output register **305_3**, respectively.

[0056] As shown in FIG. **10**, the cumulative addition circuit **402** inputs calculation results of the product calculation circuit **302** and contents of the cumulative addition register **404** for each channel on the basis of the channel

numbers of the channel register **307** and adds each channel in parallel. Thereby, a cumulative addition result for each channel can be obtained at the same time. The convolution arithmetic processing circuit **101** separately performs cumulative addition for each channel number instead of cumulatively adding the results of the product calculation.

[0057] In a specific example, an example is described where the 3×3 convolution arithmetic operation for three channels is performed at once by using the 3×3 convolution arithmetic operation circuit. However, the arithmetic operation is not limited to this, and in the embodiment, it is possible to perform an N×N convolution arithmetic operation for N channels at once by using an N×N convolution arithmetic operation circuit.

Modified Example

[0058] Hereinafter, a typical modified example will be illustrated. In a description of the modified example below, the same reference numerals as those of the above embodiment may be used for portions having the same configurations and functions as those described in the above embodiment. For the descriptions of these portions, the descriptions in the above embodiment may be appropriately applied to the extent where there is no technical contradiction. Further, parts of the above embodiment and all or parts of the modified example may by appropriately and complexly applied to the extent where there is no technical contradiction.

[0059] An example where a convolution arithmetic operation is performed regardless of an arrangement of an input image will be described with reference to FIG. **11**. FIG. **11** is a block diagram for explaining another configuration example of the convolution arithmetic processing circuit.

[0060] A convolution arithmetic processing circuit **101**A has an input pattern register **308** and an input pattern generation circuit **309** in addition to the configuration of the convolution arithmetic processing circuit **101** of the embodiment. The input pattern register **308** is a register for setting how to output pixel values set in the line buffer **301** from the input pattern generation circuit **309**. The input pattern generation circuit **309** is a circuit that receives input from the line buffer **301** and generates an output according to the input pattern register **308**. The input pattern register **308** is set by the CPU **12**.

[0061] Next, an example where the integration coefficient table generation device **50** generates one type of integration coefficient table from two types of input coefficient tables and performs the convolution arithmetic operation by using the integration coefficient table will be described with reference to FIGS. **12** to **14**. FIG. **12** is a diagram for explaining a configuration of a convolution arithmetic processing circuit. FIG. **13** is a schematic diagram for explaining an operation of an integration coefficient table generation device in FIG. **12**. FIG. **13** indicates a case of generating one integration coefficient table from input coefficient tables for two channels. FIG. **14** is a schematic diagram for explaining an operation of an input pattern generation circuit in FIG. **12**. FIG. **14** indicates a case of generating a 5×5 input image formed from 3×3 input images for two channels.

[0062] The convolution arithmetic processing circuit **101**A in FIG. **12** performs two 3×3 convolution arithmetic operations at once by using a 5×5 convolution arithmetic operation circuit.

[0063] When the input coefficient tables are $C_{3,3,FLT_1}$ and $C_{3,3,FLT_2}$, and the integration coefficient table is $C_{5,5,FLT_A}$, the following equation is established.

[Formula 5]

$$C_{5,5,FLT_A} = \begin{bmatrix} FLT_1: & FLT_1: & FLT_1: & FLT_1: & FLT_1: \\ C_{0,0,FLT_1} & C_{0,1,FLT_1} & C_{0,2,FLT_1} & C_{1,0,FLT_1} & C_{1,1,FLT_1} \\ FLT_1: & FLT_1: & FLT_1: & FLT_1: & FLT_2: \\ C_{1,2,FLT_1} & C_{2,0,FLT_1} & C_{2,1,FLT_1} & C_{2,2,FLT_1} & C_{0,0,FLT_2} \\ FLT_2: & FLT_2: & FLT_2: & FLT_2: & FLT_2: \\ C_{0,1,FLT_2} & C_{0,2,FLT_2} & C_{1,0,FLT_2} & C_{1,1,FLT_2} & C_{1,2,FLT_2} \\ FLT_2: & FLT_2: & FLT_2: & 0 & 0 \\ C_{2,0,FLT_2} & C_{2,1,FLT_2} & C_{2,2,FLT_2} & & \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad [数\ 5]$$

[0064] Further, $I_{i,j}$ is defined as a pixel value (gray value) of ith row and jth column of an input image, an output of the input pattern generation circuit **309** is defined as $A_{5,5}$, and the following equation is established.

[Formula 6]

$$A_{5,5} = \begin{bmatrix} I_{i,j} & I_{i+1,j} & I_{i+2,j} & I_{i+1,j} & I_{i+1,j+1} \\ I_{i+2,j+1} & I_{i,j+2} & I_{i+1,j+2} & I_{i+2,j+2} & I_{i,j} \\ I_{i+1,j} & I_{i+2,j} & I_{i+1,j} & I_{i+1,j+1} & I_{i+2,j+1} \\ I_{i,j+2} & I_{i+1,j+2} & I_{i+2,j+2} & & \end{bmatrix} \quad [数\ 6]$$

[0065] FIG. **13** schematically shows how the integration coefficient table is generated when specifically 1=3 and w=3. The integration coefficient table generation device **50** processes two types of input coefficient tables (FLT_**1** and FLT_**2**) and outputs an integration coefficient table (FLT_A). The coefficients of the 3×3 input coefficient tables (FLT_**1** and FLT_**2**) are outputted to the 5×5 integration coefficient table.

[0066] As shown in FIG. **13**, from left to right of the first row of FLT_A, coefficients "40", "87", and "21" in the first row of FLT_**1** and coefficients "100" and "45" in the second row of FLT_**1** are arranged. From left to right of the second row of FLT_A, the right end coefficient "14" in the second

row of FLT_**1**, coefficients "9", "33" and "165" in the third row of FLT_**1**, and the left end coefficient "7" in the first row of FLT_**2** are arranged. From left to right of the third row of FLT_A, the center coefficient "45" and the right end coefficient "56" in the first row of FLT_**2** and coefficients "25", "30", and "212" in the second row of FLT_**2** are arranged. From left to right of the fourth row of FLT_A, coefficients "120", "2", and "27" in the third row of FLT_**2** are arranged. No coefficient is arranged in the right two positions in the fourth row of FLT_A and all positions in the fifth row of FLT_A, and these positions become blank spaces. In addition to the coefficients, the channel numbers are also stored in the integration coefficient table (FLT_A). As shown in FIG. **14**, the input pattern generation circuit **309** outputs pixel values in a form where the 3×3 pixels acquired from the line buffer **301** are repeated based on the 5×5 input pattern of the input pattern register **308**. The input pattern defines an arrangement order of the first 3×3 pixels (1-1, 1-2, . . . , and 1-9) and the second 3×3 pixels (2-1, 2-2, . . . , and 2-9).

[0067] As shown in FIG. **14**, from left to right of the first row of the output of the input pattern generation circuit **309**, pixel values "76", "120", and "32" in the first row of an input image and pixel values "77" and "130" in the second row of the input image are arranged. From left to right of the second row, the right end pixel value "34" in the second row of the input image, pixel values "130", "100", and "54" in the third row of the input image, and the left end pixel value "78" in the first row of the second input image are arranged. From left to right of the third row, the center pixel value "120" and the right end pixel value "32" in the first row of the input image and pixel values "77", "130", and "34" in the second row of the input image are arranged. From left to right of the fourth row, pixel values "130", "110", and "54" in the third row of the input image are arranged. No pixel is arranged in the right two positions in the fourth row and all positions in the fifth row, and these positions become blank spaces. Thereby, the arrangement of the pixels of the output of the input pattern generation circuit **309** is the same as the arrangement of the coefficients of the integration coefficient table (FLT_A).

[0068] It is possible to read the pixels stored in the line buffer **301** in an arbitrary form by adding the input pattern register **308** and the input pattern generation circuit **309**.

[0069] In the same manner as in the embodiment, the convolution arithmetic processing circuit **101**A reads the coefficients and the channel numbers from the integration coefficient table and sets their values in the coefficient register **306** and the channel register **307**, respectively.

[0070] In a manner similar to that in the embodiment, the product calculation circuit **302** calculates products of the values set in the coefficient register **306** and the values generated by the input pattern generation circuit **309** all at once by a 5×5 product calculation circuit. The channel selection circuit **303** sets outputs of the cumulative addition circuit **402** to the output registers **305** on the basis of the values (channel numbers) set in the channel register **307**. The output control register **304** sets a correspondence between the output registers **305** and the channel register **307**. In this example, the first output register **305_1** corresponds to FLT_**1** and the second output register **305_2** corresponds to FLT_**2**. The convolution arithmetic processing circuit **101**A outputs results of performing a convolution arithmetic operation by using the coefficients of FLT_**1** and FLT_**2** to the first output register **305_1** and the second output register **305_2**, respectively. In the same manner as in the embodiment, the cumulative addition circuit **402** inputs calculation results of the product calculation circuit and contents of the cumulative addition register for each channel on the basis of the channel numbers of the channel register **307** and adds each channel in parallel. Thereby, a cumulative addition result for each channel can be obtained at the same time.

[0071] In the modified example, it is possible to perform an arbitrary convolution arithmetic operation by using the input pattern register **308** and the integration coefficient table generation device **50** in combination with each other. Therefore, even when a 5×5 convolution arithmetic operation circuit is used, it is possible to efficiently perform a 3×3 or 1×1 convolution arithmetic operation.

[0072] In a specific example, an example is described where two 3×3 convolution arithmetic operations are performed at once by using the 5×5 convolution arithmetic operation circuit. However, the arithmetic operation is not limited to this, and in the modified example, it is possible to perform L number of M×M convolution arithmetic operations at once by using an N×N convolution arithmetic operation circuit. Here, $M<N$ and $L=INT(N^2/M^2)\geq 2$.

[0073] While the invention made by the inventors has been specifically described based on the embodiment and the modified example, it is needless to say that the present invention is not limited to the embodiment and the modified example and may be variously modified.

What is claimed is:

1. A semiconductor device comprising:

an image recognition device having a convolution arithmetic processing circuit,

wherein the convolution arithmetic processing circuit includes

a coefficient register where coefficients of an integration coefficient table are set,

a product calculation circuit that calculates products of an input image and the coefficients,

a channel register where a channel number of the integration coefficient table is set,

a channel selection circuit that selects an output destination of a cumulative addition arithmetic operation on the basis of the channel number, and

a plurality of output registers that store a result of the cumulative addition arithmetic operation,

wherein the integration coefficient table is a table where a plurality of input coefficient tables are integrated and the integration coefficient table has a size of N×N, and

wherein the product calculation circuit can calculate data of N×N all at once.

2. The semiconductor device according to claim **1**,

wherein the channel selection circuit includes

a cumulative addition circuit that adds the products for each channel number, and

a cumulative addition register that holds an output of the cumulative addition circuit for each channel number, and

wherein the cumulative addition circuit performs the addition for each channel number in parallel.

3. The semiconductor device according to claim **2**,

wherein the channel selection circuit further includes a selector that outputs the output of the cumulative

addition circuit to anyone of the output registers on the basis of the channel number.

4. The semiconductor device according to claim **1**,

wherein in each element of the integration coefficient table, a coefficient having a largest value among coefficients in each corresponding element in the input coefficient tables is stored.

5. The semiconductor device according to claim **1**,

wherein the number of the input coefficient tables is N and each of the input coefficient tables has a size of N×N.

6. The semiconductor device according to claim **1**, further comprising:

an input pattern register; and

an input pattern generation circuit that changes an arrangement of an input image based on content of the input pattern register,

wherein the integration coefficient table is a table where L input coefficient tables are integrated, the integration coefficient table has a size of N×N, each of the L input coefficient tables has a size of M×M, and M is a value smaller than N, and

wherein the product calculation circuit can calculate data of N×N all at once.

7. The semiconductor device according to claim **6**,

wherein the channel selection circuit includes

a cumulative addition circuit that adds the products for each channel number, and

a cumulative addition register that holds an output of the cumulative addition circuit for each channel number, and

wherein the cumulative addition circuit performs the addition for each channel number in parallel.

8. The semiconductor device according to claim **7**,

wherein the channel selection circuit further includes a selector that outputs the output of the cumulative addition circuit to anyone of the output registers on the basis of the channel number.

9. The semiconductor device according to claim **6**,

wherein the input pattern generation circuit repeats an input image of M×M L times and converts the L input images of M×M into an input image of N×N.

10. An image recognition system comprising:

a semiconductor device; and

a semiconductor storage device,

wherein the semiconductor device includes

a convolution arithmetic processing circuit including a coefficient register where coefficients of an integration coefficient table are set, a product calculation circuit that calculates products of an input image read from the semiconductor storage device and the coefficients, a channel register where a channel number of the integration coefficient table is set, a channel selection circuit that selects an output destination of a cumulative addition arithmetic operation on the basis of the channel number, and a plurality of output registers that store a result of the cumulative addition arithmetic operation, and

a pooling circuit,

wherein the product calculation circuit can calculate data of N×N all at once, and

wherein the integration coefficient table is a table where a plurality of input coefficient tables are integrated and the integration coefficient table has a size of N×N.

11. The image recognition system according to claim **10**,

wherein the channel selection circuit includes

a cumulative addition circuit that adds the products for each channel number, and

a cumulative addition register that holds an output of the cumulative addition circuit for each channel number, and

wherein the cumulative addition circuit performs the addition for each channel number in parallel.

12. The image recognition system according to claim **11**,

wherein the channel selection circuit further includes a selector that outputs the output of the cumulative addition circuit to anyone of the output registers on the basis of the channel number.

13. he image recognition system according to claim **10**,

wherein in each element of the integration coefficient table, a coefficient having a largest value among coefficients in each corresponding element in the input coefficient tables is stored.

14. he image recognition system according to claim **10**,

wherein the number of the input coefficient tables is N and each of the input coefficient tables has a size of N×N.

15. The image recognition system according to claim **10**,

wherein the semiconductor device further includes

an input pattern register, and

an input pattern generation circuit that changes an arrangement of an input image based on content of the input pattern register,

wherein the integration coefficient table is a table where L input coefficient tables are integrated, the integration coefficient table has a size of N×N, each of the L input coefficient tables has a size of M×M, and M is a value smaller than N, and

wherein the product calculation circuit can calculate data of N×N all at once.

16. The image recognition system according to claim **15**,

wherein the channel selection circuit includes

a cumulative addition circuit that adds the products for each channel number, and

a cumulative addition register that holds an output of the cumulative addition circuit for each channel number, and

wherein the cumulative addition circuit performs the addition for each channel number in parallel.

17. The image recognition system according to claim **16**,

wherein the channel selection circuit further includes a selector that outputs the output of the cumulative addition circuit to anyone of the output registers on the basis of the channel number.

18. The image recognition system according to claim **15**,

wherein the input pattern generation circuit repeats an input image of M×M L times and converts the L input images of M×M into an input image of N×N.

* * * * *