

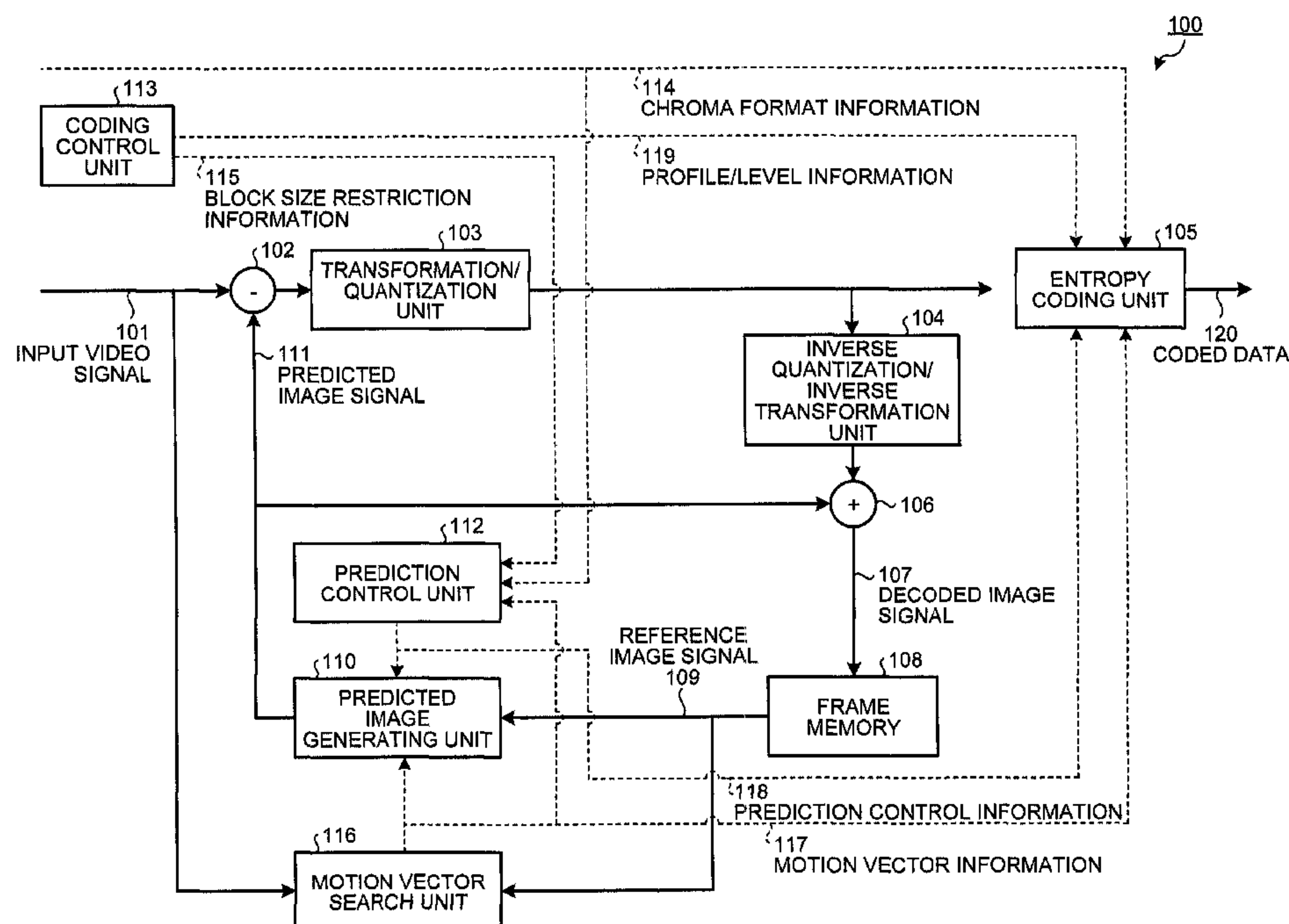


(86) Date de dépôt PCT/PCT Filing Date: 2011/11/08
(87) Date publication PCT/PCT Publication Date: 2013/05/16
(45) Date de délivrance/Issue Date: 2017/08/01
(85) Entrée phase nationale/National Entry: 2014/02/28
(86) N° demande PCT/PCT Application No.: JP 2011/075725
(87) N° publication PCT/PCT Publication No.: 2013/069095

(51) Cl.Int./Int.Cl. *H04N 19/59* (2014.01),
H04N 11/04 (2006.01), *H04N 19/176* (2014.01),
H04N 19/186 (2014.01), *H04N 19/577* (2014.01),
H04N 19/61 (2014.01)
(72) Inventeurs/Inventors:
CHUJOH, TAKESHI, JP;
YAMAKAGE, TOMOO, JP
(73) Propriétaire/Owner:
KABUSHIKI KAISHA TOSHIBA, JP
(74) Agent: MARKS & CLERK

(54) Titre : PROCEDE DE CODAGE D'IMAGE, PROCEDE DE DECODAGE D'IMAGE, DISPOSITIF DE CODAGE D'IMAGE ET DISPOSITIF DE DECODAGE D'IMAGE

(54) Title: IMAGE CODING METHOD, IMAGE DECODING METHOD, IMAGE CODING APPARATUS, AND IMAGE DECODING APPARATUS



(57) Abrégé/Abstract:

The present invention aims to reduce a memory bandwidth during image coding. An image coding method includes an acquiring step and a generating step. The acquiring step includes acquiring a reference image. The generating step includes generating a

(57) **Abrégé(suite)/Abstract(continued):**

predicted image having resolution larger than that of the reference image, by performing an interpolation to the acquired reference image according to a motion vector for each of the luminance component and the color difference component. The generating step includes generating the predicted image having the color difference component without performing a specific interpolation that means the interpolation processing in which a number of pixels to be accessed in the reference image is large, when a size of a block designated as a unit of the interpolation is smaller than a predetermined threshold value.

ABSTRACT

The present invention aims to reduce a memory bandwidth during image coding. An image coding method includes an acquiring step and a generating step. The
5 acquiring step includes acquiring a reference image. The generating step includes generating a predicted image having resolution larger than that of the reference image, by performing an interpolation to the acquired reference image according to a motion vector for each of the
10 luminance component and the color difference component. The generating step includes generating the predicted image having the color difference component without performing a specific interpolation that means the interpolation processing in which a number of pixels to be accessed in
15 the reference image is large, when a size of a block designated as a unit of the interpolation is smaller than a predetermined threshold value.

DESCRIPTION

IMAGE CODING METHOD, IMAGE DECODING METHOD, IMAGE CODING
APPARATUS, AND IMAGE DECODING APPARATUS

Field

5 [0001] Embodiments of the present invention relate to an image coding method, an image decoding method, an image coding apparatus, and an image decoding apparatus.

Background

[0002] In a technique of video coding and decoding, a motion-compensated interpolation for each block is generally executed. An image signal to be referred is stored in an external memory; therefore, when the video coding and decoding is implemented by hardware, there might be constraints on the amount of read data. Accordingly, 10 when an amount of access to the memory increases, a so-called memory bandwidth, which is a bottleneck in the coding and decoding operations, becomes a problem.

[0003] In the motion-compensated interpolation for each block, an interpolation filtering process using FIR (Finite Impulse Response) filter in the horizontal direction and in the vertical direction is executed. In the interpolation filtering process, a pixel outside the block has to be accessed. When the number of pixels outside the block increases, the memory bandwidth per pixel also increases.

25 [0004] Conventionally, the memory bandwidth per pixel has been reduced by applying an interpolation filter with a short tap length to a block with a small size by which the ratio of the accessed pixels outside the block relatively increases.

30 Citation List

Patent Literature

[0005] Patent Literature 1: Japanese Patent No. 4120301

Summary

[0006] However, in the conventional art, the memory bandwidth cannot appropriately be reduced. For example, in the case of the chroma format, such as 4 : 2 : 0 or 4 : 2 : 2, in which the number of samples of the pixel for a color difference (color difference component) is smaller than the number of samples of the pixel for luminance (luminance component), and the resolution is low, the interpolation has to be executed in such a manner that the color difference is enlarged more with the luminance being defined as a reference. Therefore, when the filter with taps longer than two taps is used for the interpolation for the color difference, the process for the color difference signal cannot be restricted, even if the process is changed per luminance block.

[0007] An image coding method of an embodiment includes an acquiring step and a generating step. The acquiring step includes acquiring a reference image. The generating step includes generating a predicted image having resolution larger than that of the reference image by performing, for each of the luminance component and the color difference component, an interpolation on the acquired reference image according to a motion vector. Further, the generating step includes generating the predicted image for the color difference component without performing a specific interpolation that means the interpolation in which a number of pixels to be accessed in the reference image is large, when a size of a block, which is designated as a unit of the interpolation, is smaller than a predetermined threshold value.

According to an aspect of the present invention there is provided an image decoding method of decoding a target image including a luminance component and color difference components, the method comprising:

2a

acquiring a reference image;

selecting a bi-prediction or a uni-prediction based on
a prediction mode that is explicitly or implicitly
designated, the bi-prediction generating a predicted image
5 by using two reference images, the uni-prediction
generating the predicted image by using a single reference
image;

selecting the bi-prediction for the target image if a
prediction mode designating the bi-prediction, wherein a
10 block size of the target image is equal to or smaller than
a first threshold value,

after selecting the bi-prediction, determining whether
the block size of the target image is equal to or smaller
than the first threshold value,

15 if the block size of the target image is determined to
be equal to or smaller than the first threshold value
changing the bi-prediction to the uni-prediction, and

generating the predicted image by interpolating the
luminance component and the color difference components in
20 the reference image according to a motion vector.

According to another aspect of the present invention
there is provided an image decoding apparatus that decodes
a target image including a luminance component and color
difference components, the apparatus comprising:

25 a generating unit configured to generate a predicted
image by interpolating the luminance component and the
color difference components in a reference image according
to a motion vector; and

a coding unit configured to code coefficient
30 information acquired from a prediction error indicating a
difference between the predicted image and to the target
image, wherein

the generating unit selects a bi-prediction or a uni-

prediction based on a prediction mode that is explicitly or implicitly designated, the bi-prediction generating a predicted image by using two reference images, the uni-prediction generating the predicted image by using a single
5 reference image;

selects the bi-prediction for the target image if a prediction mode designating the bi-prediction, wherein a block size of the target image is equal to or smaller than a first threshold value, and
10 after selecting the bi-prediction, determines whether the block size of the target image is equal to or smaller than the first threshold value,
if the block size of the target image is determined to be equal to or smaller than the first threshold value,
15 changing the bi-prediction to the uni-prediction, and performs the interpolation on the reference image to generate the predicted image according to the motion vector.

According to a further aspect of the present invention there is provided an image coding method of coding a target
20 image including a luminance component and color difference components, the method comprising:

acquiring a reference image;
selecting a bi-directional prediction or a uni-directional prediction based on a prediction mode that is
25 explicitly or implicitly designated;
if a size of a block, which is designated as a unit of an interpolation, is equal to or smaller than a first threshold value, and if the bi-directional prediction is selected, changing the bi-directional prediction to the
30 uni-directional prediction, and
generating a predicted image by interpolating the luminance component and the color difference components in the reference image according to a motion vector.

According to a further aspect of the present invention there is provided an image coding apparatus that codes a target image including a luminance component and color difference components, the apparatus comprising:

5 a generating unit configured to generate a predicted image by interpolating the luminance component and the color difference components in a reference image according to a motion vector; and

10 a coding unit configured to code coefficient information acquired from a prediction error indicating a difference between the predicted image and to the target image, wherein

15 the generating unit selects a bi-directional prediction or a uni-directional prediction based on a prediction mode that is explicitly or implicitly designated, if a size of a block, which is designated as a unit of an interpolation, is equal to or smaller than a first threshold value, and if the bi-directional prediction is selected, changing the bi-directional prediction to the
20 uni-directional prediction, and performs the interpolation on the reference image to generate the predicted image according to the motion vector.

Brief Description of Drawings

25 [0008] FIG. 1 is a block diagram illustrating an image

coding apparatus according to an embodiment of the present invention.

FIG. 2 is a view illustrating one example of chroma format information.

5 FIG. 3 is a view of a motion vector with a color difference signal in 4 : 2 : 0 format.

FIG. 4 is a view of a motion vector with a luminance signal in 4 : 2 : 0 format.

10 FIG. 5 is a view of a motion vector with a color difference signal in 4 : 2 : 2 format.

FIG. 6 is a view of a motion vector with a luminance signal in 4 : 2 : 2 format.

FIG. 7 is a view illustrating an example of pixels that are accessed in 4 : 2 : 0 format.

15 FIG. 8 is a view illustrating an example of pixels that are accessed in 4 : 2 : 0 format.

FIG. 9 is a view illustrating an example of pixels that are accessed in 4 : 2 : 2 format.

20 FIG. 10 is a block diagram illustrating an image decoding apparatus corresponding to the image coding apparatus.

FIG. 11 is a block diagram illustrating a predicted image generating unit.

25 FIG. 12 is a flowchart of control according to the embodiment.

FIG. 13 is a flowchart of a process of reducing a memory bandwidth.

FIG. 14 is a flowchart of a process of reducing a memory bandwidth.

30 FIG. 15 is a flowchart of a process of reducing a memory bandwidth.

FIG. 16 is a diagram illustrating a hardware configuration of the apparatus according to the embodiment.

Description of Embodiments

[0009] Preferable embodiments of the image coding method, the image decoding method, the image coding apparatus, and the image decoding apparatus according to the present invention will be described below in detail with reference to the accompanying drawings.

[0010] The image coding apparatus and the image decoding apparatus according to the present embodiment make a control by referring to chroma format information so that a position of a pixel indicated by a motion vector in a block having a size smaller than a predetermined size is not subject to an interpolation for a color difference, thereby reducing a memory bandwidth.

[0011] FIG. 1 is a block diagram illustrating one example of a configuration of an image coding apparatus 100 according to the present embodiment. As illustrated in FIG. 1, the image coding apparatus 100 includes a subtraction unit 102, a transformation/quantization unit 103, an inverse quantization/inverse transformation unit 104, an entropy coding unit 105, an addition unit 106, a frame memory 108, a predicted image generating unit 110, a prediction control unit 112, a coding control unit 113, and a motion vector search unit 116.

[0012] The image coding apparatus 100 generates coded data 120 from input video signal 101. For example, the input video signal 101 is input to the image coding apparatus 100 in units of frames. The input video signal 101 is divided into a block that is a macroblock.

[0013] The subtraction unit 102 outputs a prediction error signal that is a difference between a predicted image signal 111 generated by the predicted image generating unit 110 and the input video signal 101.

[0014] The transformation/quantization unit 103 executes

a quantization after executing an orthogonal transformation on prediction error signal with a discrete cosine transformation (DCT), thereby generating quantized transform coefficient information. The quantized transform coefficient information is divided into two. One of the divided information is input to the entropy coding unit 105. The other one is input to the inverse quantization/inverse transformation unit 104.

[0015] The inverse quantization/inverse transformation unit 104 executes the inverse quantization and inverse transformation on the quantized transform coefficient information as the process inverse to the processing executed by the transformation/quantization unit 103, thereby reproducing the prediction error signal.

[0016] The addition unit 106 adds the prediction error signal and the predicted image signal. According to this process, a decoded image signal 107 is generated. The decoded image signal 107 is input to the frame memory 108.

[0017] The frame memory 108 is a memory unit that stores therein a reference image signal. The frame memory 108 executes a filtering process or the other process on the decoded image signal 107, and then, determines whether the decoded image signal 107 is stored or not for allowing the decoded image signal 107 to become the reference image signal 109 input to the predicted image generating unit 110. The reference image signal 109 is input to the predicted image generating unit 110 and to the motion vector search unit 116.

[0018] The motion vector search unit 116 generates motion vector information 117 from the input video signal 101 and the reference image signal 109. The motion vector information 117 is input to the predicted image generating unit 110, and also transmitted to the entropy coding unit

105.

[0019] The predicted image generating unit 110 generates the predicted image signal 111 from the reference image signal 109, the prediction control information 118, and the motion vector information 117.

[0020] The coding control unit 113 inputs block size restriction information 115 to the prediction control unit 112, and transmits profile/level information 119 to the entropy coding unit 105.

[0021] The profile/level information 119 includes profile information indicating a combination of coding tool groups, and level information that is restriction information of the image coding apparatus according to the processing power of the image decoding apparatus. The level information indicates a restriction combination of a maximum number of macroblocks per hour, the maximum number of macroblocks per frame, the maximum search range of vector, and the number of vectors in two consecutive macroblocks.

[0022] For example, H.264 specifies profile information such as a base line profile, a main profile, and high profile. H.264 also specifies 16 level information.

[0023] In the present embodiment, parameters are specified using the profile/level information. The parameters includes a parameter specifying as to whether the memory bandwidth reducing method is applied or not, the restriction value of the block size (block size restriction information 115), and a restriction method. These parameters may be specified by using information other than the profile/level information.

[0024] The block size restriction information 115 is information specifying a threshold value (the restriction value of the block size) used for the determination of the

block size. For example, the coding control unit 113 sets different block size restriction information 115 according to the profile/level information. The block size restriction information 115 may be included in the
5 profile/level information.

[0025] The prediction control unit 112 controls the predicted image generation executed by the predicted image generating unit 110 according to the block size restriction information 115 input from the coding control unit 113,
10 chroma format information 114 of the input video signal 101, and the motion vector information 117 input from the motion vector search unit 116 (the detail will be described later). The prediction control unit 112 generates the prediction control information 118 used for the control of the
15 predicted image generation. The prediction control information 118 is input to the predicted image generating unit 110, and also transmitted to the entropy coding unit 105.

[0026] The entropy coding unit 105 performs an entropy
20 coding on the coding information to generate the coded data 120 according to a prescribed syntax. The coding information includes, for example, the quantized transform coefficient information input from the transformation/quantization unit 103, the chroma format
25 information 114 of the input video signal, the motion vector information 117 input from the motion vector search unit 116, the prediction control information 118 input from the prediction control unit 112, and the profile/level information 119 input from the coding control unit 113.

30 [0027] Here, the chroma format information 114 will be described. The chroma format information 114 is information indicating a chroma format of the input video signal 101. FIG. 2 is a view illustrating one example of

the chroma format information 114. FIG. 2 illustrates an example in which chroma_format_idc used in H.264 is used as the chroma format information 114.

[0028] chroma_format_idc = 0 indicates a monochrome
5 format only with luminance. chroma_format_idc = 1
indicates 4 : 2 : 0 format in which the color difference is
sampled at half horizontally and vertically with respect to
the luminance. chroma_format_idc = 2 indicates 4 : 2 : 2
10 format in which the color difference is sampled at half
only horizontally with respect to the luminance.
chroma_format_idc = 3 indicates 4 : 4 : 4 format in which
the luminance and the color difference have the same pixel
number.

[0029] The horizontal size of the prediction block of
15 the luminance signal is defined as nPSW, and the vertical
size is defined as nPSH. In 4 : 2 : 0 format, the
horizontal size of the blocks of the color difference
signals Cb and Cr is nPSW/2, while the vertical size is
nPSH/2. In 4 : 2 : 2 format, the horizontal size of the
20 blocks of the color difference signals Cb and Cr is nPSW/2,
while the vertical size is nPSH. In 4 : 4 : 4 format, the
horizontal size of the blocks of the color difference
signals Cb and Cr is nPSW, while the vertical size is nPSH.

[0030] Next, the relationship between the chroma format
25 and the interpolation will be described.

[0031] FIG. 3 is a view illustrating the position of the
motion vector in an interpolation image with 1/8-pel
accuracy of the color difference signal in 4 : 2 : 0 format.
"B" is a position of an integer pixel of the color
30 difference signal, which is the position of the motion
vector that does not need the interpolation. White
portions indicate the position of the motion vector that
needs a one-dimensional interpolation for the color

difference signal only horizontally or only vertically.
Light shaded portions indicate the position of the motion
vector that needs a two-dimensional interpolation for
performing the interpolation to the color difference signal
5 both horizontally and vertically.

[0032] FIG. 4 is a view illustrating the position of the
motion vector in an interpolation image with 1/4-pel
accuracy of the luminance signal in 4 : 2 : 0 format. "A"
is the position of the integer pixel of the luminance
10 signal, which is the position of the motion vector that
does not need the interpolation. White portions with "A"
indicate the position of the motion vector that does not
need the interpolation for both the luminance signal and
the color difference signal. Light shaded portions with
15 "A" indicate the position of the motion vector that does
not need the interpolation for the luminance signal but
needs the interpolation for the color difference signal.

[0033] The white portions without "A" indicate the
position of the motion vector that needs the one-
20 dimensional interpolation for the luminance signal and the
color difference signal only horizontally or only
vertically. The light shaded portions without "A" indicate
the position of the motion vector that needs the two-
dimensional interpolation in which the interpolation
25 processing is performed horizontally and vertically for the
luminance signal and the color difference signal. Dark
shaded portions indicate the position of the motion vector
that needs the one-dimensional interpolation only
horizontally or only vertically for the luminance signal,
30 and needs the two-dimensional interpolation in which the
interpolation is executed horizontally and vertically for
the color difference signal.

[0034] FIG. 5 is a view illustrating the position of the

motion vector in an interpolation image with 1/4-pel accuracy of the color difference signal in the horizontal direction, and with 1/8-pel accuracy of the color difference signal in the vertical direction in 4 : 2 : 2 format. "B" is the position of the integer pixel of the color difference signal, which is the position of the motion vector that does not need the interpolation. White portions indicate the position of the motion vector that needs the one-dimensional interpolation for the color difference signal only horizontally or only vertically. Light shaded portions indicate the position of the motion vector that needs the two-dimensional interpolation for performing the interpolation to the color difference signal horizontally and vertically.

[0035] FIG. 6 is a view illustrating the position of the motion vector in an interpolation image with 1/4-pel accuracy of the luminance signal in 4 : 2 : 2 format. "A" is the position of the integer pixel of the luminance signal, which is the position of the motion vector that does not need the interpolation for the luminance signal. White portions with "A" indicate the position of the motion vector that does not need the interpolation for both the luminance signal and the color difference signal. Light shaded portions with "A" indicate the position of the motion vector that does not need the interpolation for the luminance signal but needs the interpolation for the color difference signal.

[0036] The white portions without "A" indicate the position of the motion vector that needs the one-dimensional interpolation for the luminance signal and the color difference signal only horizontally or only vertically. The light shaded portions without "A" indicate the position of the motion vector that needs the two-

dimensional interpolation in which the interpolation is performed horizontally and vertically for the luminance signal and the color difference signal. Dark shaded portions indicate the position of the motion vector that needs the one-dimensional interpolation only horizontally for the luminance signal, and needs the two-dimensional interpolation in which the interpolation is executed horizontally and vertically for the color difference signal.

[0037] Next, the relationship between the chroma format and the pixel to be accessed in the interpolation will be described.

[0038] FIGS. 7 and 8 are views illustrating one example of a pixel that is accessed upon generating the interpolation image on the block basis in 4 : 2 : 0 format.

[0039] FIG. 7 illustrates the maximum number of pixels that have to be accessed upon generating the interpolation image of 4x4 pixel block for the luminance signal with an 8-tap interpolation filter. In the two-dimensional interpolation, three outside pixels on the left and above the pixel block as well as four outside pixels on the right and under the pixel block have to be accessed for generating the interpolation image with 4x4 pixel block. Specifically, 11x11 pixels have to be accessed as a whole. The number of the outside pixels to be accessed depends upon the tap length. Therefore, when the interpolation filter with the same tap is used, the number of accesses per pixel increases more for a smaller block.

[0040] FIG. 8 illustrates the maximum number of pixels that have to be accessed upon generating the interpolation image of 2x2 pixel block, corresponding to 4x4 pixel block for the luminance signal, for the color difference signal with a four-tap interpolation filter. In the two-

dimensional interpolation, one outside pixel on the left and above the pixel block as well as two outside pixels on the right and under the pixel block have to be accessed for generating the interpolation image with 2x2 pixel block.

5 Specifically, 5x5 pixels have to be accessed as a whole.

[0041] FIG. 9 is a view illustrating one example of a pixel that is accessed upon generating the interpolation image on the block basis in 4 : 2 : 2 format. The maximum number of pixels that have to be accessed upon generating
10 the interpolation image of 4x4 pixel block for the luminance signal with a four-tap interpolation filter is the same as the case in FIG. 7, so that the redundant description will not be made.

[0042] FIG. 9 illustrates the maximum number of pixels
15 that have to be accessed upon generating the interpolation image of 4x2 pixel block, corresponding to 4x4 pixel block for the luminance signal, for the color difference signal with a four-tap interpolation filter. In the two-dimensional interpolation, one outside pixel on the left
20 and above the pixel block as well as two outside pixels on the right and under the pixel block have to be accessed for generating the interpolation image with 2x2 pixel block. Specifically, 5x7 pixels have to be accessed as a whole.

[0043] As illustrated in FIGS. 3 to 6, the necessity of
25 the interpolation is different depending upon the chroma format and the motion vector. Which is needed out of the one-dimensional interpolation and the two-dimensional interpolation is different depending upon the chroma format and the motion vector. As illustrated in FIGS. 7 to 9, the
30 number of pixels to be accessed is different depending upon the chroma format.

[0044] In the present embodiment, by referring to the

chroma format and the motion vector, the predicted image generation is controlled so that a specific interpolation in which the number of pixels to be accessed in the reference image (reference image signal 109) is large is not executed. The specific interpolation is an interpolation using bi-directional prediction and two-dimensional interpolation. The interpolation in the bi-directional prediction may be defined as the specific interpolation. The specific method for controlling the predicted image generation so as not to execute the specific interpolation will be described later.

[0045] FIG. 10 is a block diagram illustrating an example of a configuration of an image decoding apparatus 300 corresponding to the image coding apparatus 100. The image decoding apparatus 300 includes an entropy decoding unit 302, an inverse quantization/inverse transformation unit 303, an addition unit 304, a frame memory 306, and the predicted image generating unit 110.

[0046] The image decoding apparatus 300 generates a reproduced video signal 307 from coded data 301.

[0047] The entropy decoding unit 302 performs an entropy decoding on the coded data 301 in accordance with a prescribed syntax. The entropy decoding unit 302 decodes the coded data 301 to acquire quantized transform coefficient information, prediction control information 311, motion vector information 312, and profile/level information 313. The decoded quantized transform coefficient information is input to the inverse quantization/inverse transformation unit 303. The decoded prediction control information 311, the motion vector information 312, and the profile/level information 313 are input to the predicted image generating unit 110.

[0048] The quantized transform coefficient information,

the prediction control information 311, the motion vector information 312, and the profile/level information 313 correspond respectively to the quantized transform coefficient information, the prediction control information 118, the motion vector information 117, and the profile/level information 119, which are coded by the image coding apparatus 100 in FIG. 1.

[0049] The inverse quantization/inverse transformation unit 303 executes inverse quantization and inverse orthogonal transformation on the quantized transform coefficient information, thereby reproducing the prediction error signal.

[0050] The addition unit 304 adds the prediction error signal and the predicted image signal 310 to generate a decoded image signal 305. The decoded image signal 305 is input to the frame memory 306.

[0051] The frame memory 306 executes the filtering process on the decoded image signal 305, and outputs the resultant as the reproduced video signal 307. The frame memory 306 determines whether the decoded image signal 305, which has undergone the filtering process, is stored or not, based upon the prediction control information 311. The stored decoded image signal 305 is input to the predicted image generating unit 110 as a reference image signal 308.

[0052] The predicted image generating unit 110 generates the predicted image signal 310 by using the reference image signal 308, the prediction control information 311, and the motion vector information 312.

[0053] FIG. 11 is a block diagram illustrating an example of a configuration of the predicted image generating unit 110 mounted to the image coding apparatus 100 and the image decoding apparatus 300. The predicted image generating unit 110 includes a switch 201, a bi-

directional prediction unit 202, a uni-directional prediction unit 203, and an intra-prediction unit 204. The predicted image generating unit 110 generates the predicted image signal 111 from the reference image signal 109, the prediction control information 118, and the motion vector information 117.

[0054] The prediction control information 118 includes information (prediction mode) for designating which one of the bi-directional prediction unit 202, the uni-directional prediction unit 203, and the intra-prediction unit 204 is used, for example. The switch 201 makes a changeover for selecting any one of the bi-directional prediction unit 202, the uni-directional prediction unit 203, and the intra-prediction unit 204 by referring to this information.

[0055] The reference image signal 109 is input to any one of the bi-directional prediction unit 202, the uni-directional prediction unit 203, and the intra-prediction unit 204, which is selected by the switch 201.

[0056] When the bi-directional prediction unit 202 is selected, the bi-directional prediction unit 202 generates a motion compensation image signal by using the reference image signal 109 and the motion vector information 117 from plural reference frames, and generates the predicted image signal 111 based upon the bi-directional prediction. The bi-directional prediction unit 202 is selected not only in the case where the prediction mode is explicitly designated as the bi-directional prediction as the coded data but also in the case where the bi-directional prediction is not explicitly designated by the coded data such as a skip mode, a direct mode, and merge mode, but the operation of the bi-directional prediction is implicitly designated by semantics.

[0057] When the uni-directional prediction unit 203 is

selected, the uni-directional prediction unit 203 generates the motion compensation image signal by using the reference image signal 109 and the motion vector information 117 from a single reference frame, and generates the predicted image signal 111. The uni-directional prediction unit 203 is selected not only in the case where the prediction mode is explicitly designated as the uni-directional prediction as the coded data but also in the case where the uni-directional prediction is not explicitly designated by the coded data such as the skip mode, the direct mode, and the merge mode, but the operation of the uni-directional prediction is implicitly designated by semantics.

[0058] When the intra-prediction unit 204 is selected, the intra-prediction unit 204 generates the predicted image signal 111 by using the reference image signal 109 in a screen.

[0059] Next, the control for reducing the memory bandwidth by the image coding apparatus 100 thus configured according to the present embodiment will be described with reference to FIG. 12. FIG. 12 is a flowchart illustrating an overall flow of the control in the present embodiment.

[0060] The coding control unit 113 sets a restriction value (nLPSW, nLPSH) of the block size according to the profile/level information 119 (step S101). nLPSW is the restriction value of the predicted block size of luminance in the horizontal direction. nLPSH is the restriction value of the predicted block size of luminance in the vertical direction.

[0061] When the profile information indicates a specific profile (e.g., high profile of H.264), or when the level information indicates a specific level (e.g., a certain level or higher level), for example, the coding control unit 113 sets the predetermined restriction value (nLPSW,

nLPSH) of the block size. The coding control unit 113 may be configured to set stepwise the restriction value of the block size according to the profile information and the level information.

5 [0062] It is supposed below that a variable RW is a motion vector accuracy in the horizontal direction, expressed by $1/RW$ -pel accuracy. It is also supposed that a variable RH is a motion vector accuracy in the vertical direction, expressed by $1/RH$ -pel accuracy. Initial values
10 of the variable RW and the variable RH are defined as the motion vector accuracy of luminance. A value of a power of two is generally used for RW and RH.

[0063] The prediction control unit 112 determines whether the chroma format information (chroma_format_idc)
15 114 is 1 or not (step S102). In the case of chroma_format_idc = 1 (step S102: Yes), the prediction control unit 112 doubles the values of RW and RH (step S103). This is because chroma_format_idc = 1 means 4 : 2 : 0 format in which the color difference is sampled at half
20 horizontally and vertically with respect to luminance.

[0064] In the case where chroma_format_idc = 1 is not established (step S102: No), the prediction control unit 112 determines whether the chroma format information (chroma_format_idc) 114 is 2 or not (step S104). In the
25 case of chroma_format_idc = 2 (step S104: Yes), the prediction control unit 112 doubles the value of RW (step S105). This is because chroma_format_idc = 2 means 4 : 2 : 2 format in which the color difference is sampled at half only horizontally with respect to luminance.

30 [0065] When chroma_format_idc assumes other values (step S104: No), the values of RW and RH are not changed.

[0066] Next, the prediction control unit 112 calculates a variable L indicating whether the memory bandwidth is

restricted or not (step S106). The variable L assuming "true" means that the method of reducing the memory bandwidth is applied, and the variable L assuming "false" means that the method is not applied.

5 [0067] When the prediction is the bi-directional prediction, the prediction block is small, and two motion vectors are fractional accuracy in the color difference, as described above for example, the memory bandwidth to be accessed per pixel increases. Therefore, the prediction control unit 112 calculates the variable L according to the
10 following equation (1).

$$\begin{aligned}
 L = & (\text{PredMode} == \text{PredBi}) \ \&\& \\
 & (\text{nPSW} \leq \text{nLPSW}) \ \&\& (\text{nPSH} \leq \text{nLPSH}) \ \&\& \\
 & (\text{mvL0}[0] \ \& \ (\text{RW} - 1)) \ \&\& \\
 15 \quad & (\text{mvL0}[1] \ \& \ (\text{RH} - 1)) \ \&\& \\
 & (\text{mvL1}[0] \ \& \ (\text{RW} - 1)) \ \&\& \\
 & (\text{mvL1}[1] \ \& \ (\text{RH} - 1)); \qquad \qquad \qquad (1)
 \end{aligned}$$

[0068] The value of the motion vector in the horizontal direction in the list 0 of the block to be processed is
20 defined as mvL0[0], and the value in the vertical direction is defined as mvL0[1]. The value of the motion vector in the horizontal direction in the list 1 is defined as mvL1[0], and the value in the vertical direction is defined as mvL1[1]. PredMode indicates the prediction mode.
25 PredBi indicates the bi-directional prediction. In the description below, the prediction modes of the uni-directional prediction using the motion vectors in the list 0 and in the list 1 are represented as PredL0 and PredL1, respectively.

30 [0069] An example of the equation (1) means a case in which the prediction mode PredMode is PredBi, i.e., the bi-directional prediction unit 202 is selected. (nPSW ≤

nLPSW) && (nPSH $\leq$ nLPSH) && means the condition in which the prediction block size is equal to or smaller than the block size restriction information. In (mvL0[0] & (RW - 1)) &&, (mvL0[1] & (RH - 1)) &&, (mvL1[0] & (RW - 1)) &&, and (mvL1[1] & (RH - 1)), it is checked whether the two motion vectors L0 and L1 are not subjected to the two-dimensional interpolation for the color difference, i.e., whether the lower bit of the motion vector expresses the accuracy after the decimal point. "&" means a bit operator according to the notation in the C language, and expresses bitwise OR.

[0070] The conditional equation for calculating the variable L is not limited to the equation (1). For example, it may independently be determined even for the prediction modes (PredL0, PredL1) other than PredBi as in an equation (2).

```

L = (nPSW  $\leq$  nLPSW) && (nPSH  $\leq$  nLPSH) &&
  ((PredMode == PredBi) &&
    (mvL0[0] & (RW - 1)) &&
    (mvL0[1] & (RH - 1)) &&
    (mvL1[0] & (RW - 1)) &&
    (mvL1[1] & (RH - 1))) ||
  ((PredMode == PredL0) &&
    (mvL0[0] & (RW - 1)) &&
    (mvL0[1] & (RH - 1)) ||
    ((PredMode == PredL1) &&
      (mvL1[0] & (RW - 1)) &&
      (mvL1[1] & (RH - 1))))) ;      (2)

```

[0071] As in an equation (3), the restriction value (nLPSW1, nLPSH1) of the block size for the uni-directional prediction (PredL0 or PredL1) may separately be set. Specifically, the block size restricted in the uni-directional prediction and the block size restricted in the

20

bi-directional prediction may be different from each other.

```

    L = ((PredMode == PredBi) &&
        (nPSW ≤ nLPSW) && (nPSH ≤ nLPSH) &&
        (mvL0[0] & (RW - 1)) &&
5      (mvL0[1] & (RH - 1)) &&
        (mvL1[0] & (RW - 1)) &&
        (mvL1[1] & (RH - 1))) ||
        ((nPSW ≤ nLPSW1) &&
        (nPSH ≤ nLPSH1)) ||
10     ((PredMode == PredL0) &&
        (mvL0[0] & (RW - 1)) &&
        (mvL0[1] & (RH - 1)) &&
        ((PredMode == PredL1) &&
        (mvL1[0] & (RW - 1)) &&
15     (mvL1[1] & (RH - 1))));           (3)

```

[0072] When the prediction block size is equal to or smaller than the restriction value (nLPSW, nLPSH) of the block size as in an equation (4), two motion vectors may be restricted to access only to the integer pixel in the color difference during the bi-directional prediction.

```

    L = ((PredMode == PredBi) &&
        (nPSW ≤ nLPSW) && (nPSH ≤ nLPSH) &&
        !((mvL0[0] & (RW - 1) == 0) &&
        (mvL0[1] & (RH - 1) == 0) &&
25     (mvL1[0] & (RW - 1) == 0) &&
        (mvL1[1] & (RH - 1) == 0)));           (4)

```

[0073] Whether the value of the motion vector is restricted, or under what condition the value of the motion vector is restricted is distinguished by the profile/level information 119.

[0074] Next, the specific method of reducing the memory bandwidth will be described. FIG. 13 is a flowchart

illustrating one example of the process of reducing the memory bandwidth. FIG. 13 illustrates one example of a method of restricting the value of the motion vector, as the method of reducing the memory bandwidth.

5 [0075] The prediction control unit 112 determines whether the variable L is "true" or not (step S201). If the variable L is "true" (step S201: Yes), the prediction control unit 112 transforms the values L0 and L1 of two motion vectors as in an equation (5) (step S202).

10 [0076] $mvL0[0] = ((mvL0[0] + (RW \gg 1)) / RW) \times RW;$

$mvL0[1] = ((mvL0[1] + (RH \gg 1)) / RH) \times RH;$

$mvL1[0] = ((mvL1[0] + (RW \gg 1)) / RW) \times RW;$

$mvL1[1] = ((mvL1[1] + (RH \gg 1)) / RH) \times RH; \quad (5)$

[0077] ">>" indicates an arithmetic right shift according to the notation in the C language. "/" indicates a division in integer arithmetic. "x" indicates a multiplication in integer arithmetic. The bit corresponding to the interpolation accuracy of the color difference signal of two motion vectors L0 and L1 are rounded by the equation (5) to become 0. With this process, the two-dimensional interpolation is not executed, whereby the reduction in the memory bandwidth can be achieved.

[0078] The general rounding method is described here. However, the other method may be used. For example, rounding down method, rounding up method, and a method of rounding to the nearest even number may be employed.

[0079] The motion vector information 117 whose value is changed is coded in the entropy coding unit 105, and output as the coded data. The method in FIG. 13 is for controlling the motion vector information 117 by restricting the value of the motion vector, so as not to generate the coded data by which the memory bandwidth

increases.

[0080] Alternatively, instead of coding the motion vector information 117 whose value is changed in the entropy coding unit 105, the motion vector information 117 before the change may be coded by the entropy coding unit. In this case, the predicted image generating unit 110 in the image decoding apparatus 300 determines whether the method of reducing the memory bandwidth is applied or not in the process same as that in FIG. 12. When it is applied, the predicted image generating unit 110 in the image decoding apparatus 300 restricts the motion vector by the same manner as in FIG. 13.

[0081] The method of transforming the value of the motion vector is not limited to the method of rounding the value corresponding to the interpolation accuracy of the color difference as in the equation (4). The value may be rounded separately for the luminance and the color difference. Specifically, during the interpolation for the luminance, the value corresponding to the interpolation accuracy of the luminance may be rounded, while the value corresponding to the interpolation accuracy of the color difference may be rounded during the interpolation processing for the color difference. This method is for not generating the predicted image that increases the memory bandwidth, when the image coding apparatus 100 and the image decoding apparatus 300 are configured in advance in a manner to execute the same operation.

[0082] FIG. 14 is a flowchart illustrating another example of the process of reducing the memory bandwidth. FIG. 14 illustrates another example of the method of restricting the value of the motion vector.

[0083] In this example, the prediction control unit 112 and the predicted image generating unit 110 calculate cost

for selecting the prediction mode, the predicted block size, and the motion vector. They preferentially select the prediction mode, the predicted block size, and the motion vector, which are small in cost, whereby the optimum
5 combination can be selected.

[0084] A variable MV_Cost indicating the cost for the motion vector is calculated by using a sum of absolute distance (SAD) of predicted residual errors, a code amount of the motion vector information (MV_Code), and a Lagrange
10 multiplier (λ) calculated from the quantized information as in an equation (6).

$$\text{MV_Cost} = \text{SAD} + \lambda \times \text{MV_Code} \quad (6)$$

[0085] If the variable L is "true" (step S301: Yes), the prediction control unit 112 substitutes the predetermined
15 maximum value MaxValue into the variable MV_Cost indicating the cost for the motion vector (step S302). With this process, the prediction control unit 112 controls not to select the motion vector having the large memory bandwidth (step S301).

20 [0086] In the method in FIG. 14, the value of the motion vector is restricted to control the motion vector information 117, so as not to generate the coded data by which the memory bandwidth increases, as in FIG. 13.

[0087] FIG. 15 is a flowchart illustrating another
25 example of the method of reducing the memory bandwidth. FIG. 15 illustrates a method of controlling the prediction mode of the color difference, as another method of reducing the memory bandwidth.

[0088] If the variable L is "true" (step S401), only the
30 prediction mode PredMode of color is forcibly rewritten to the uni-directional prediction PredL0 (step S402). With this process, the case of the bi-directional prediction with the color difference signal using large memory

bandwidth can be restricted.

[0089] The prediction mode in which the prediction mode is forcibly rewritten may be the uni-directional prediction PredL1. What prediction mode is restricted is determined
5 according to the profile/level information 119.

[0090] As described above, according to the present embodiment, the memory bandwidth upon generating the motion-compensated interpolation image during the image coding and image decoding can be reduced.

10 [0091] Next, a hardware configuration of the apparatus (the image coding apparatus, and the image decoding apparatus) according to the present embodiment will be described with reference to FIG. 16. FIG. 16 is an explanatory view illustrating a hardware configuration of
15 the apparatus according to the present embodiment.

[0092] The apparatus according to the present embodiment includes a control device such as a CPU (Central Processing Unit) 51, a memory device such as a ROM (Read Only Memory) 52 or a RAM (Random Access Memory) 53, a communication I/F
20 54 that is connected to network to enable intercommunication, and a bus 61 that interconnects each unit.

[0093] A program executed by the apparatus according to the present embodiment is provided as preliminarily being
25 incorporated in the ROM 52.

[0094] The program may be configured to be provided, as a computer product, as being recorded as a file in an installable format or in an executable format to a computer-readable recording medium such as a CD (Compact
30 Disc)-ROM, a flexible disk (FD), a CD-R, a DVD (Digital Versatile Disk), and the like.

[0095] Further, the program may be provided in such a manner that the program is stored to a computer connected

to a network such as the Internet to allow download via the network. The program may be configured to be provided or distributed via a network such as the Internet.

[0096] The program executed by the apparatus according to the present embodiment can allow the computer to function as each unit (predicted image generating unit, and the like) described above. The CPU 51 in the computer can read the program from the computer-readable memory medium onto the main memory device, and executes the same program.

10 [0097] While certain embodiments have been described, these embodiments have been presented by way of example only, and are not intended to limit the scope of the inventions. Indeed, the novel embodiments described herein may be embodied in a variety of other forms; furthermore, various omissions, substitutions and changes in the form of the embodiments described herein may be made without departing from the scope of the invention. The accompanying claims and their equivalents are intended to cover such forms or modifications as would fall within the scope of the invention.

Reference Signs List

[0098] 100 IMAGE CODING APPARATUS
101 INPUT VIDEO SIGNAL
102 SUBTRACTION UNIT
25 103 TRANSFORMATION/QUANTIZATION UNIT
104 INVERSE QUANTIZATION/INVERSE TRANSFORMATION UNIT
105 ENTROPY CODING UNIT
106 ADDITION UNIT
107 DECODED IMAGE SIGNAL
30 108 FRAME MEMORY
109 REFERENCE IMAGE SIGNAL
110 PREDICTED IMAGE GENERATING UNIT
111 PREDICTED IMAGE SIGNAL

26

	112	PREDICTION CONTROL UNIT
	113	CODING CONTROL UNIT
	114	CHROMA FORMAT INFORMATION
	115	BLOCK SIZE RESTRICTION INFORMATION
5	116	MOTION VECTOR SEARCH UNIT
	117	MOTION VECTOR INFORMATION
	118	PREDICTION CONTROL INFORMATION
	119	PROFILE/LEVEL INFORMATION
	120	CODED DATA
10	300	IMAGE DECODING APPARATUS
	301	CODED DATA
	302	ENTROPY DECODING UNIT
	303	INVERSE QUANTIZATION/INVERSE TRANSFORMATION UNIT
	304	ADDITION UNIT
15	305	DECODED IMAGE SIGNAL
	306	FRAME MEMORY
	307	REPRODUCED VIDEO SIGNAL
	308	REFERENCE IMAGE SIGNAL
	310	PREDICTED IMAGE SIGNAL
20	311	PREDICTION CONTROL INFORMATION
	312	VECTOR INFORMATION
	313	PROFILE/LEVEL INFORMATION

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

1. An image decoding method of decoding a target image including a luminance component and color difference components, the method comprising:

acquiring a reference image;

selecting a bi-prediction or a uni-prediction based on a prediction mode that is explicitly or implicitly designated, the bi-prediction generating a predicted image by using two reference images, the uni-prediction generating the predicted image by using a single reference image;

selecting the bi-prediction for the target image if a prediction mode designating the bi-prediction, wherein a block size of the target image is equal to or smaller than a first threshold value,

after selecting the bi-prediction, determining whether the block size of the target image is equal to or smaller than the first threshold value,

if the block size of the target image is determined to be equal to or smaller than the first threshold value changing the bi-prediction to the uni-prediction, and

generating the predicted image by interpolating the luminance component and the color difference components in the reference image according to a motion vector.

2. The image decoding method according to claim 1, wherein

in the generating, if the size of the block, which is designated as a unit of the interpolation, is equal to or smaller than a second threshold value that is different from the first threshold value, the uni-prediction is not performed, transforming the motion vector, which is a fractional-pel accuracy of the reference image, into a motion vector that is an integer-pel accuracy of the reference image,

and generating the predicted image from the transformed motion vector and the reference image.

3. The image decoding method according to claim 1, wherein the predicted image is generated from two reference images if the bi-prediction is performed and generated from a single reference image if the uni-prediction is performed.

4. An image decoding apparatus that decodes a target image including a luminance component and color difference components, the apparatus comprising:

a generating unit configured to generate a predicted image by interpolating the luminance component and the color difference components in a reference image according to a motion vector; and

a coding unit configured to code coefficient information acquired from a prediction error indicating a difference between the predicted image and to the target image, wherein

the generating unit selects a bi-prediction or a uni-prediction based on a prediction mode that is explicitly or implicitly designated, the bi-prediction generating a predicted image by using two reference images, the uni-prediction generating the predicted image by using a single reference image;

selects the bi-prediction for the target image if a prediction mode designating the bi-prediction, wherein a block size of the target image is equal to or smaller than a first threshold value, and

after selecting the bi-prediction, determines whether the block size of the target image is equal to or smaller than the first threshold value,

if the block size of the target image is determined to be equal to or smaller than the first threshold value, changing the bi-prediction to the uni-prediction, and performs the

interpolation on the reference image to generate the predicted image according to the motion vector.

5. The image decoding apparatus according to claim 4, wherein

if the size of the block, which is designated as a unit of the interpolation, is equal to or smaller than a second threshold value that is different from the first threshold value, the generating unit does not perform the uni-prediction, transforming the motion vector, which is a fractional-pel accuracy of the reference image, into a motion vector that is an integer-pel accuracy of the reference image, and generating the predicted image from the transformed motion vector and the reference image.

6. The image decoding apparatus according to claim 4, wherein

the predicted image is generated from two reference images if the bi-prediction is performed and generated from a single reference image if the uni-prediction is performed.

7. The image decoding apparatus according to claim 4, wherein

the generating unit and the coding unit are implemented as a processor.

8. An image coding method of coding a target image including a luminance component and color difference components, the method comprising:

acquiring a reference image;

selecting a bi-directional prediction or a uni-directional prediction based on a prediction mode that is explicitly or implicitly designated;

if a size of a block, which is designated as a unit of an interpolation, is equal to or smaller than a first threshold

value, and if the bi-directional prediction is selected, changing the bi-directional prediction to the uni-directional prediction, and

generating a predicted image by interpolating the luminance component and the color difference components in the reference image according to a motion vector.

9. The image coding method according to claim 8, wherein

the predicted image is generated from two reference images if the bi-directional prediction is performed and generated from a single reference image if the uni-directional prediction is performed.

10. An image coding apparatus that codes a target image including a luminance component and color difference components, the apparatus comprising:

a generating unit configured to generate a predicted image by interpolating the luminance component and the color difference components in a reference image according to a motion vector; and

a coding unit configured to code coefficient information acquired from a prediction error indicating a difference between the predicted image and to the target image, wherein

the generating unit selects a bi-directional prediction or a uni-directional prediction based on a prediction mode that is explicitly or implicitly designated, if a size of a block, which is designated as a unit of an interpolation, is equal to or smaller than a first threshold value, and if the bi-directional prediction is selected, changing the bi-directional prediction to the uni-directional prediction, and performs the interpolation on the reference image to generate the predicted image according to the motion vector.

11. The image coding apparatus according to claim 10, wherein

the predicted image is generated from two reference images if the bi-directional prediction is performed and generated from a single reference image if the uni-directional prediction is performed.

12. The image coding apparatus according to claim 10, wherein the generating unit and the coding unit are implemented as a processor.

FIG.1

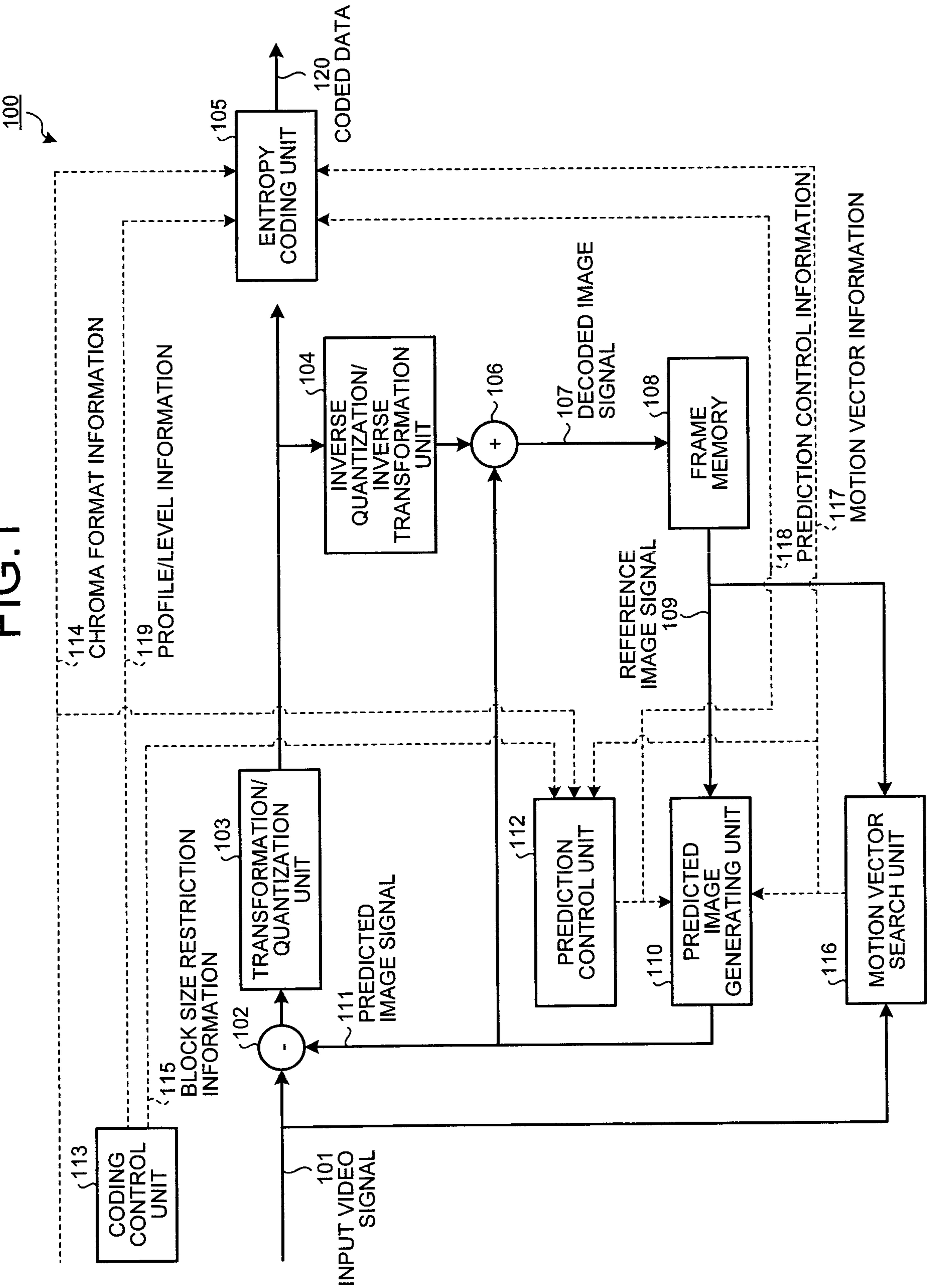
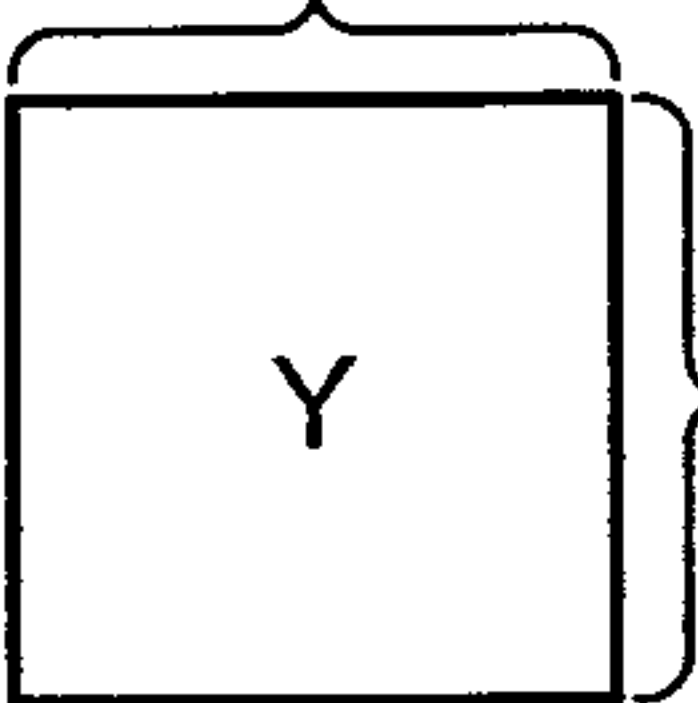
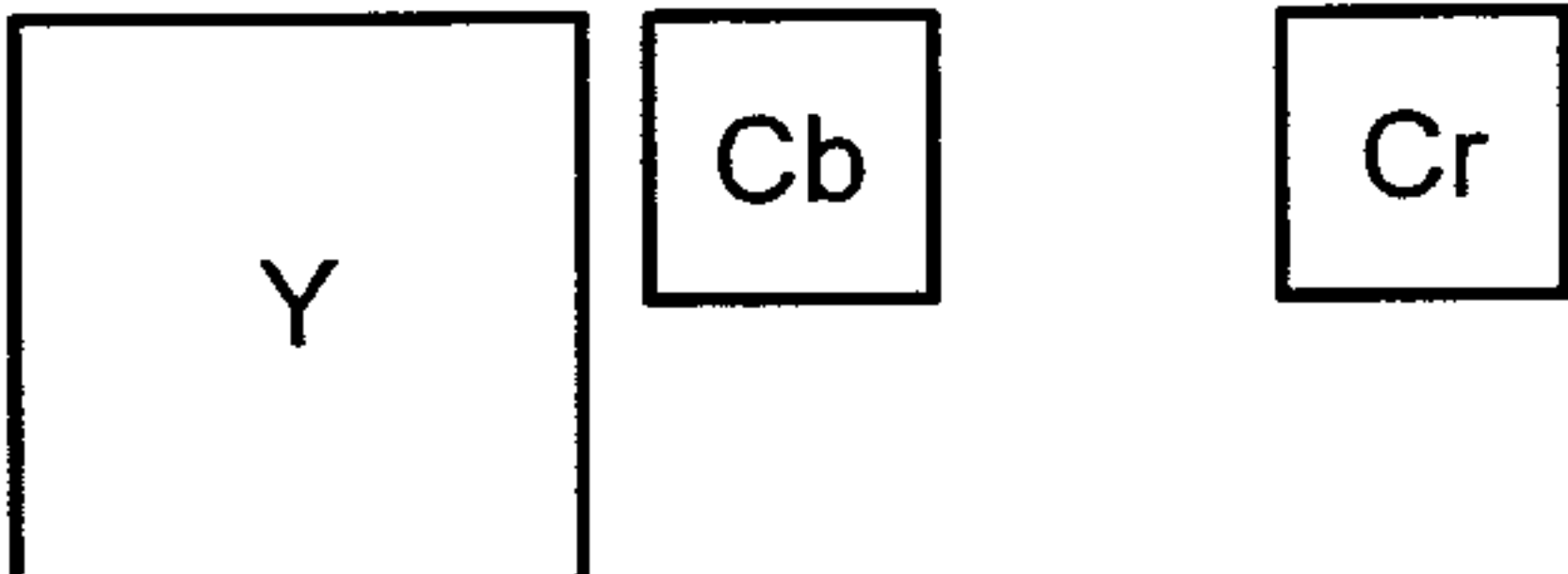
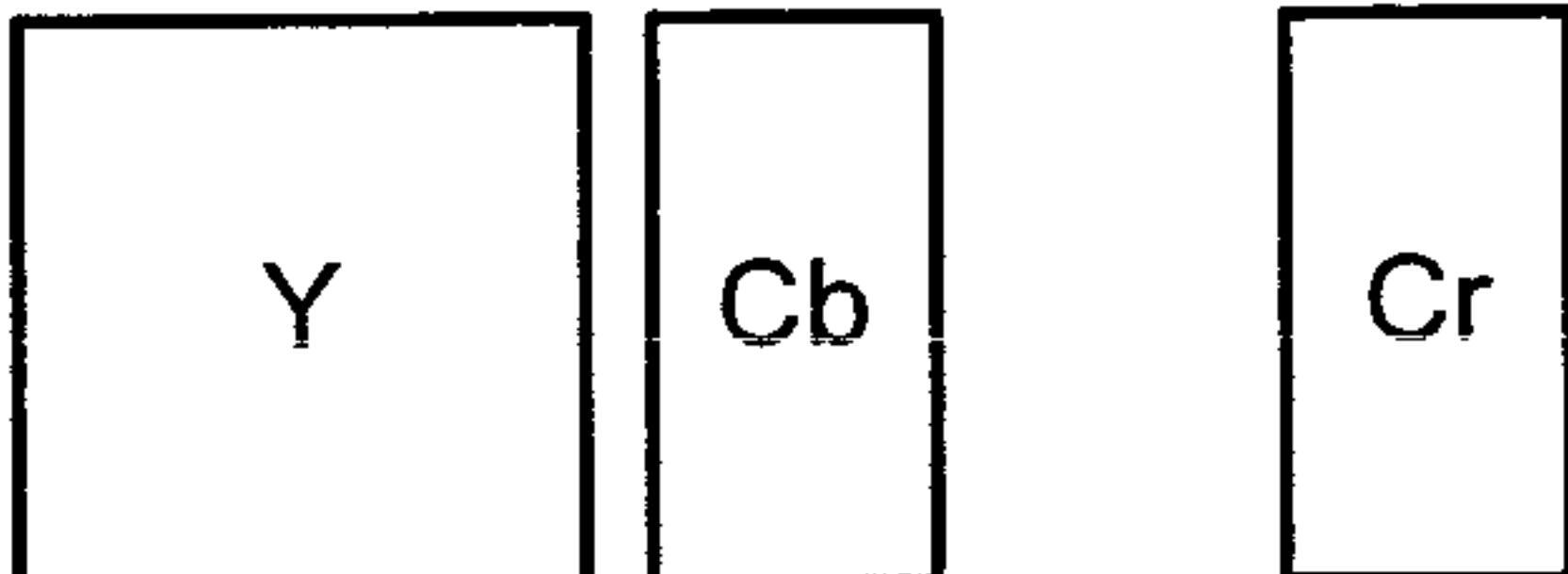
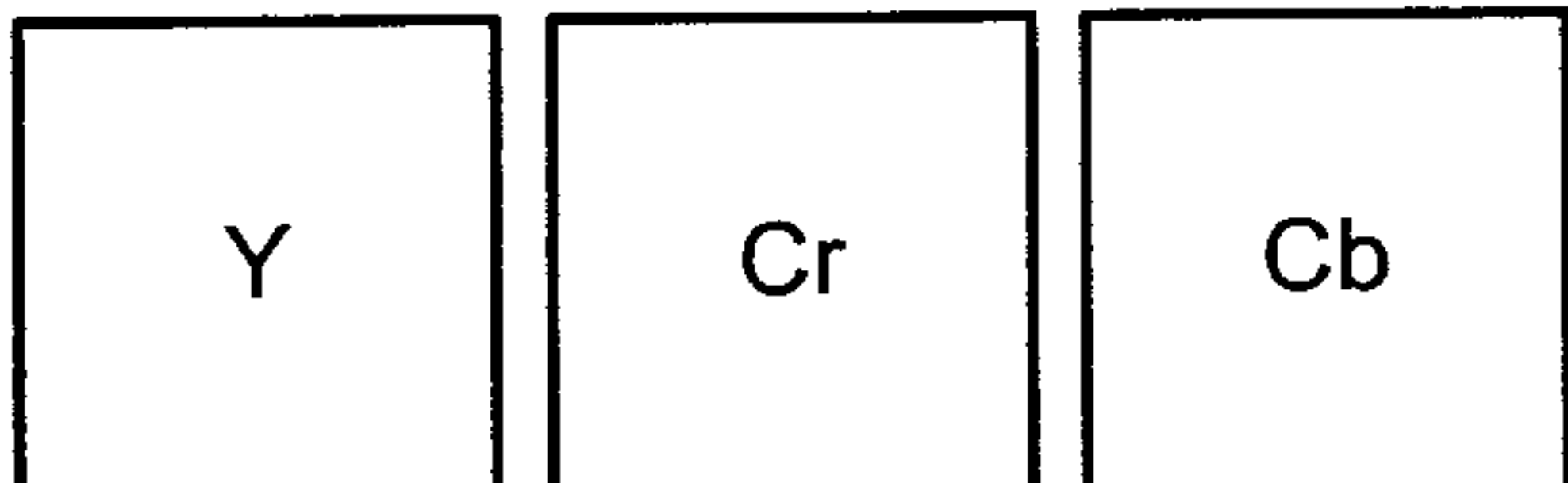


FIG.2

	MONO-CHROME	chroma_format_idc=0
	4:2:0	chroma_format_idc=1
	4:2:2	chroma_format_idc=2
	4:4:4	chroma_format_idc=3

3/13

FIG.3

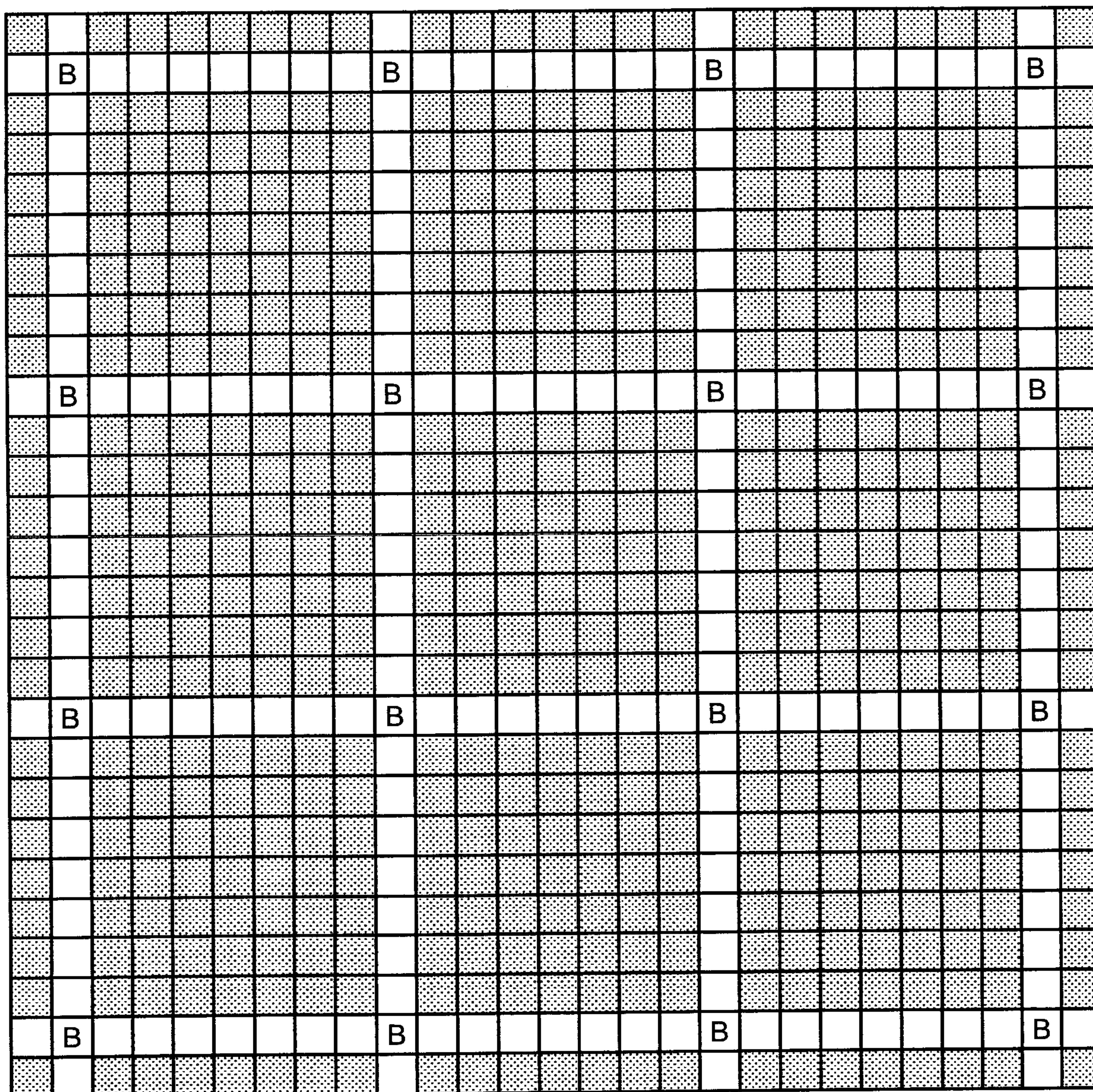


FIG.4

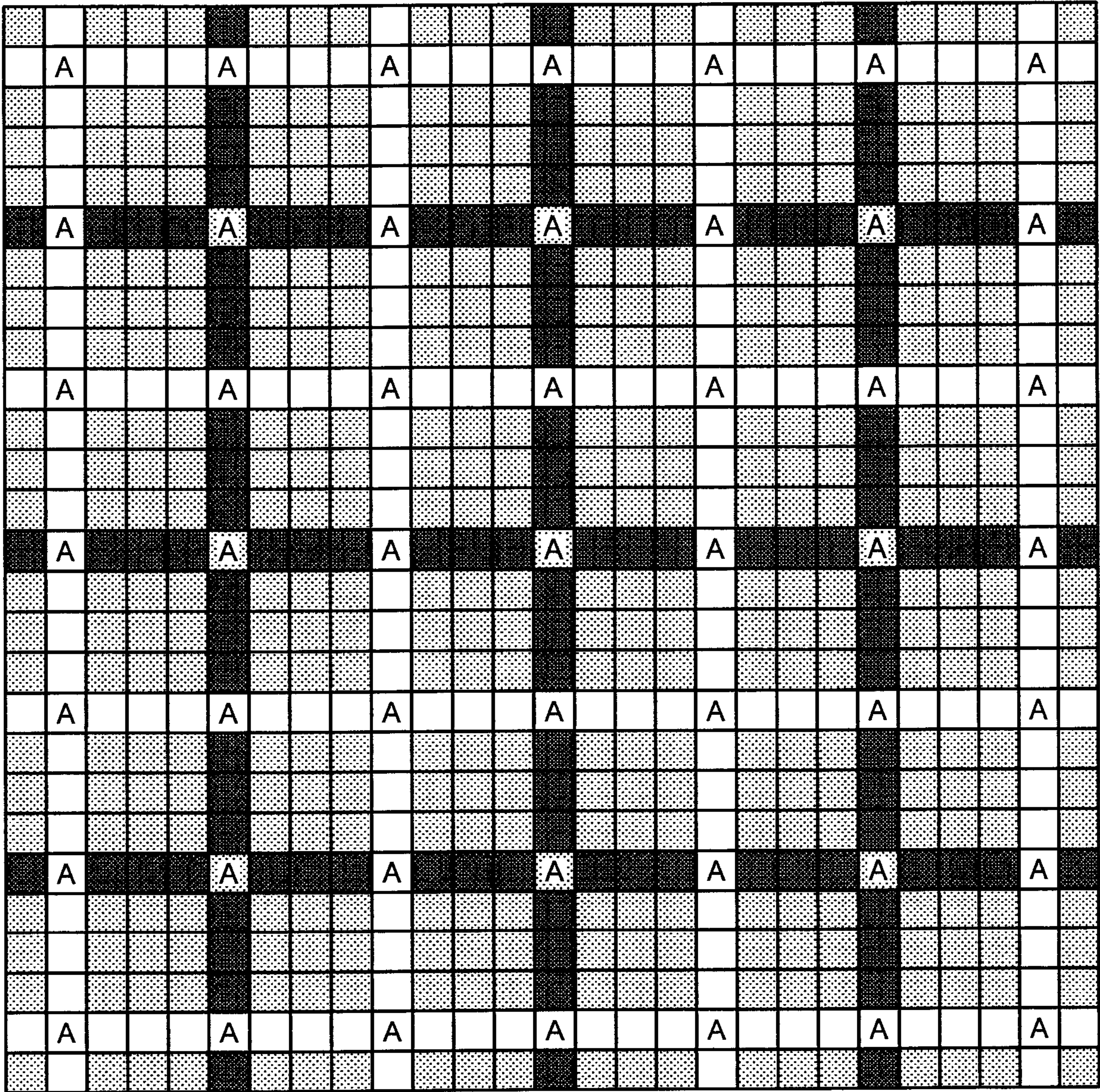


FIG.5

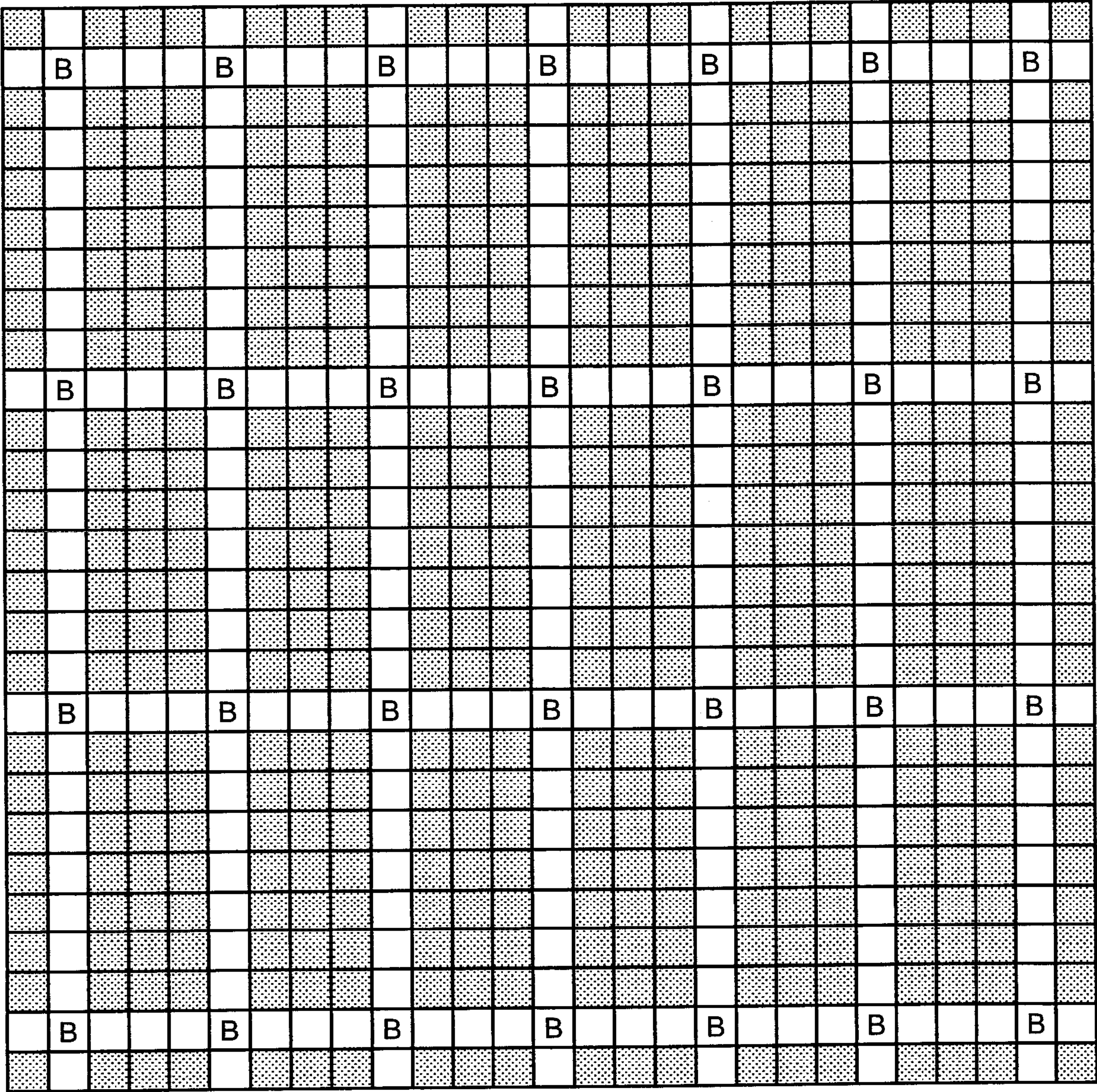
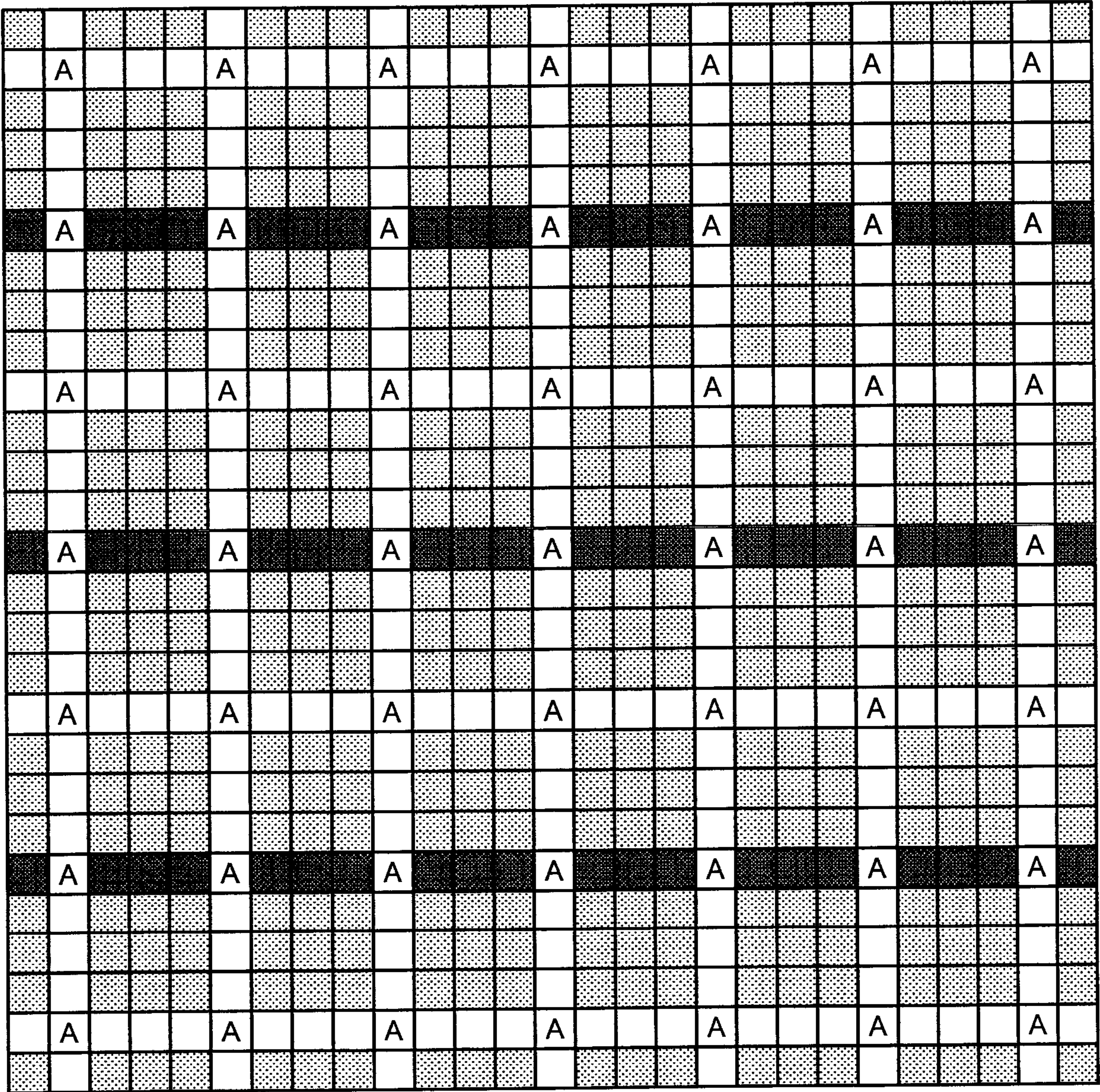


FIG.6



7/13

FIG.7

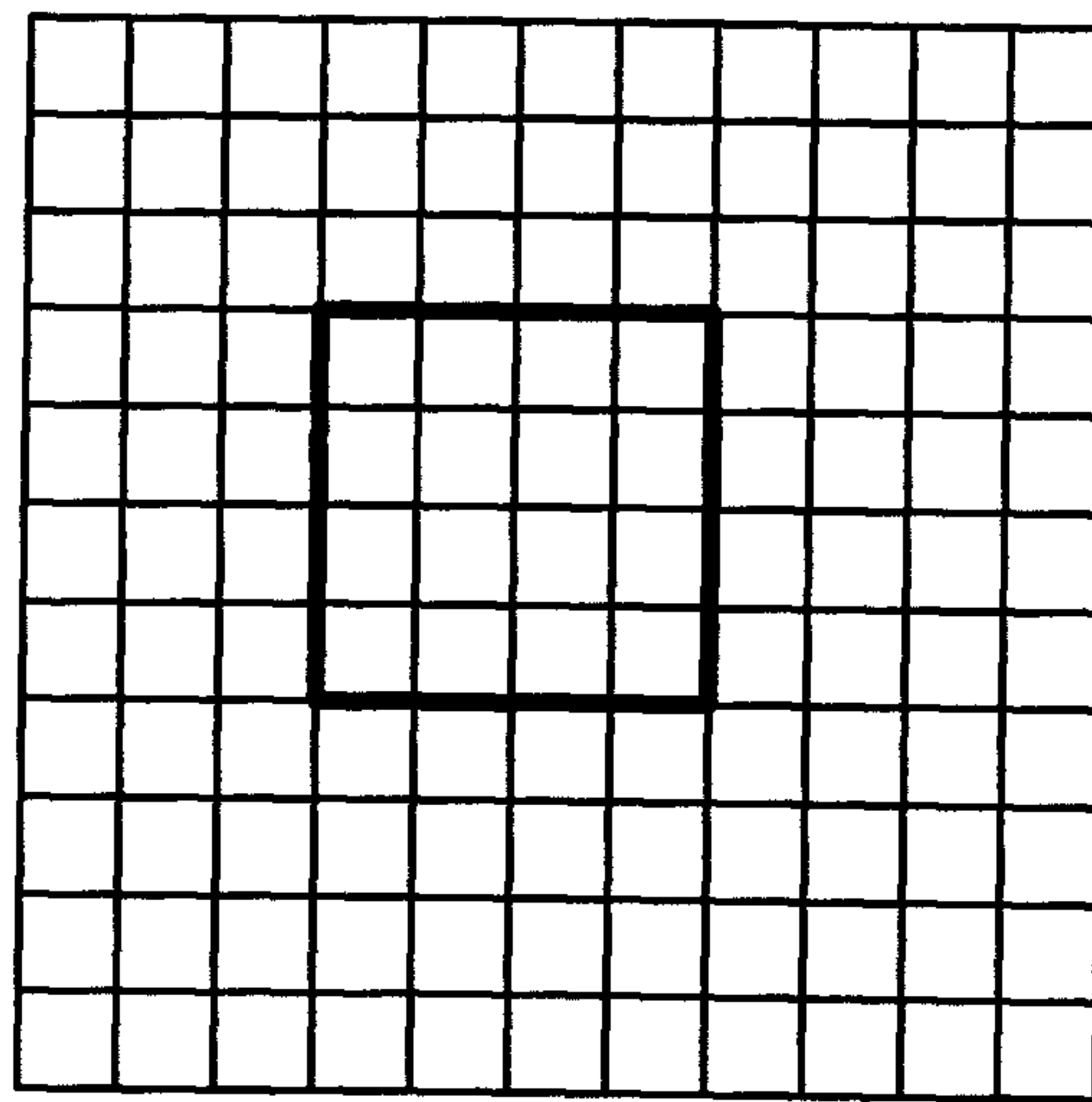
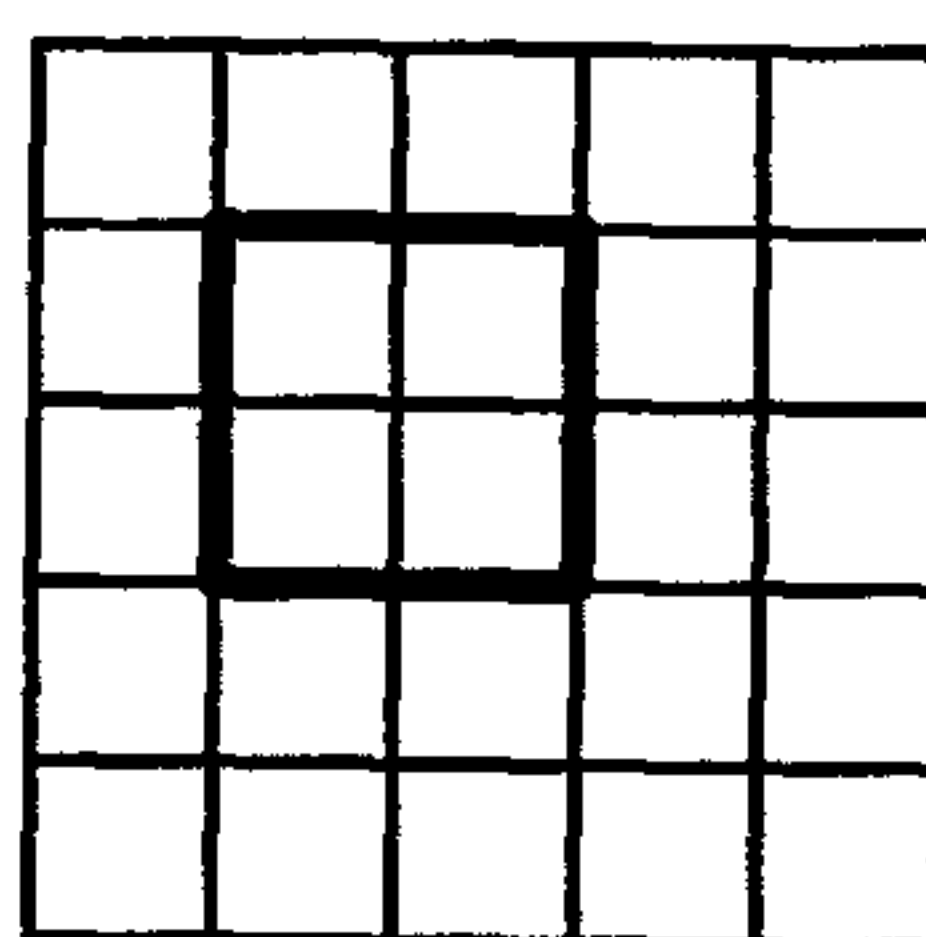
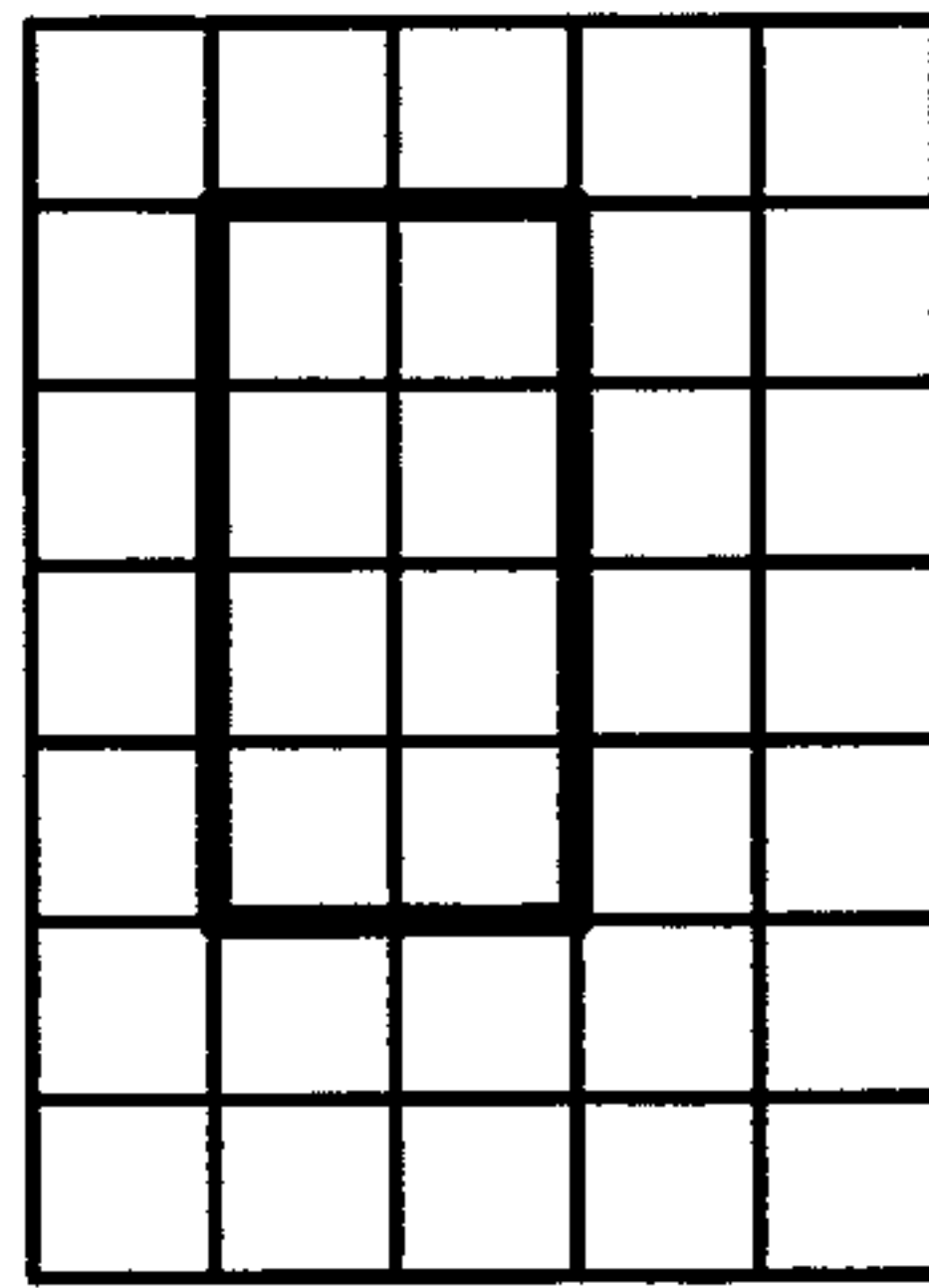


FIG.8



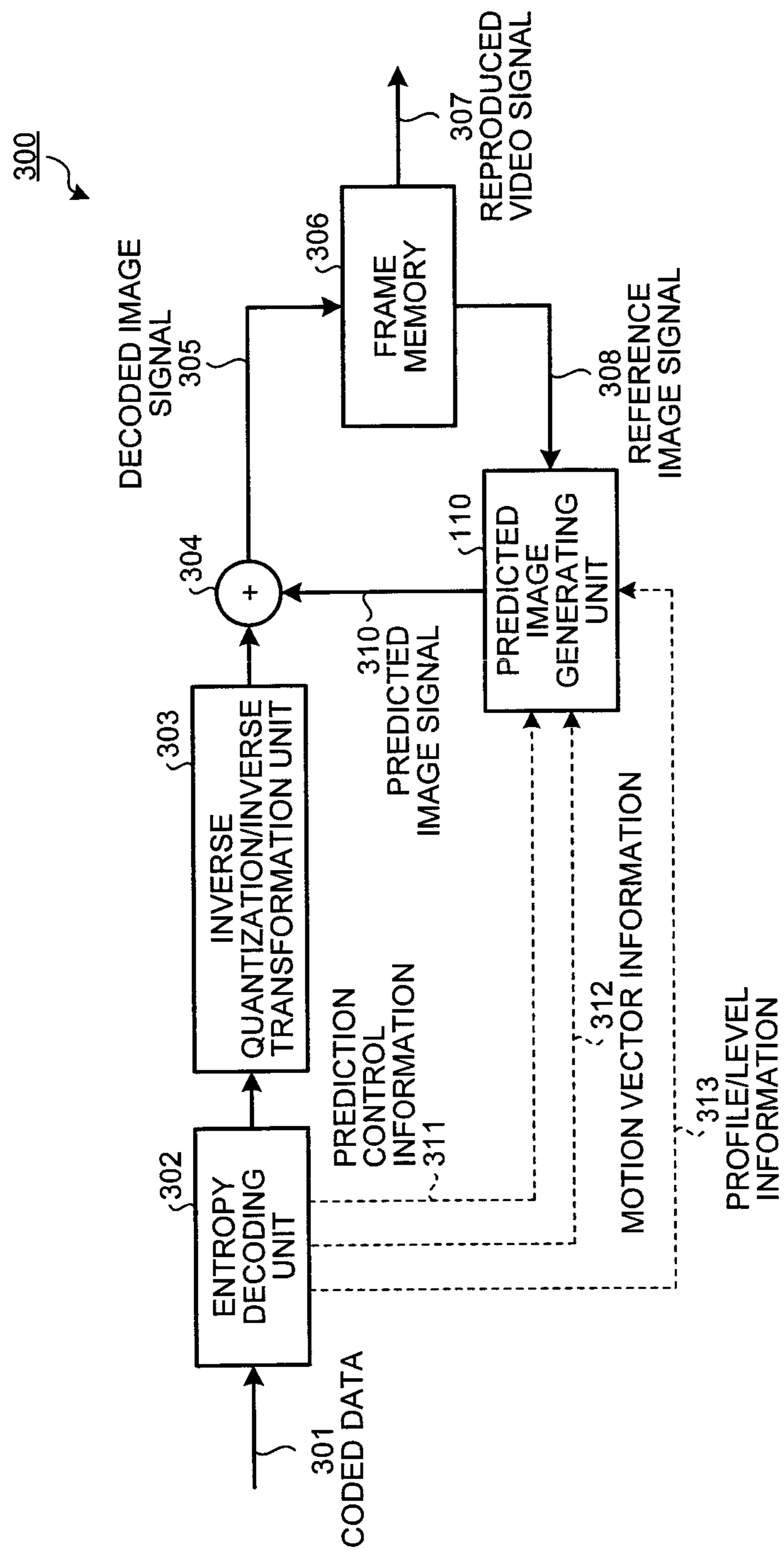
8/13

FIG.9



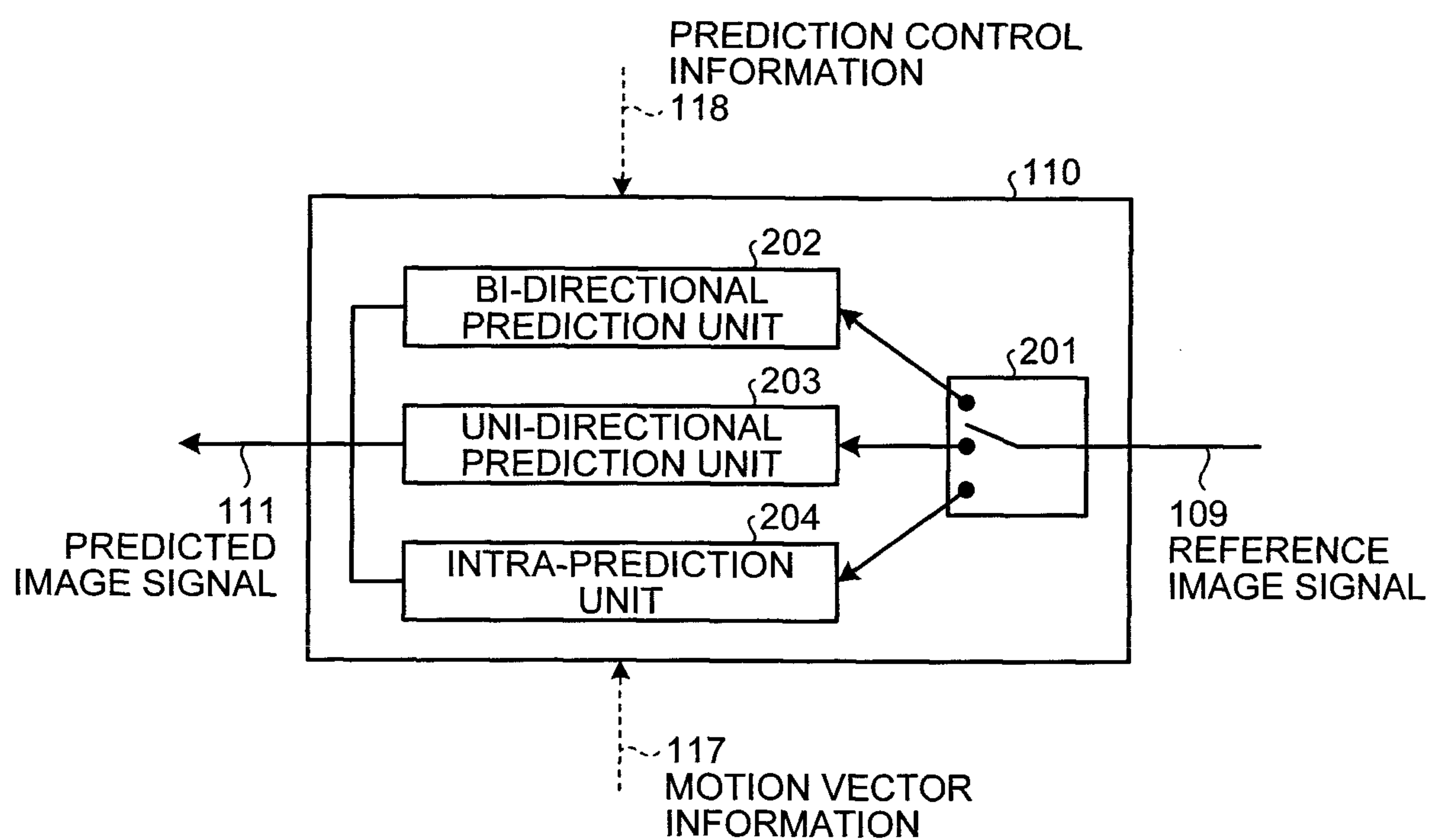
9/13

FIG. 10



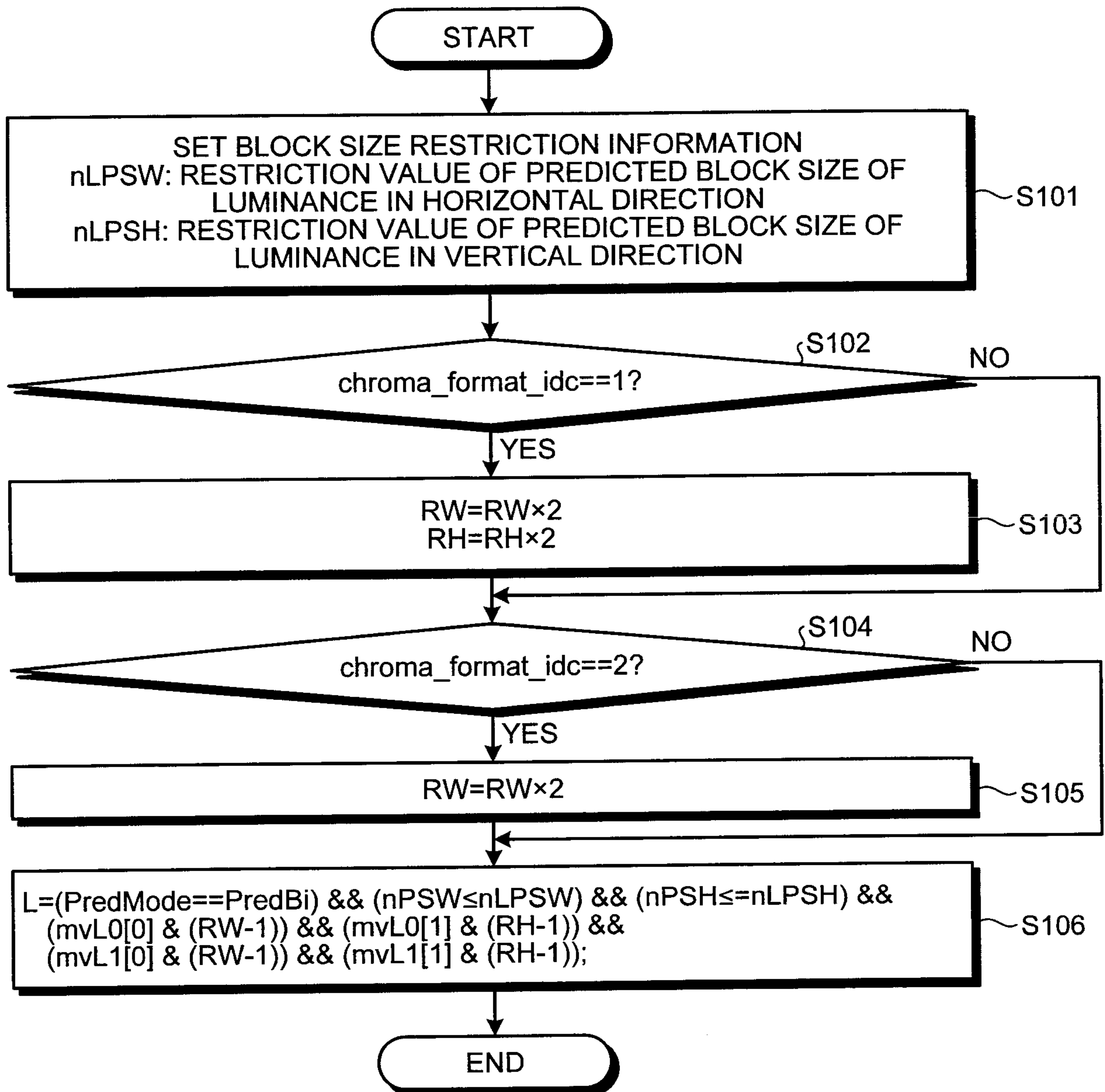
10/13

FIG.11



11/13

FIG.12



12/13

FIG.13

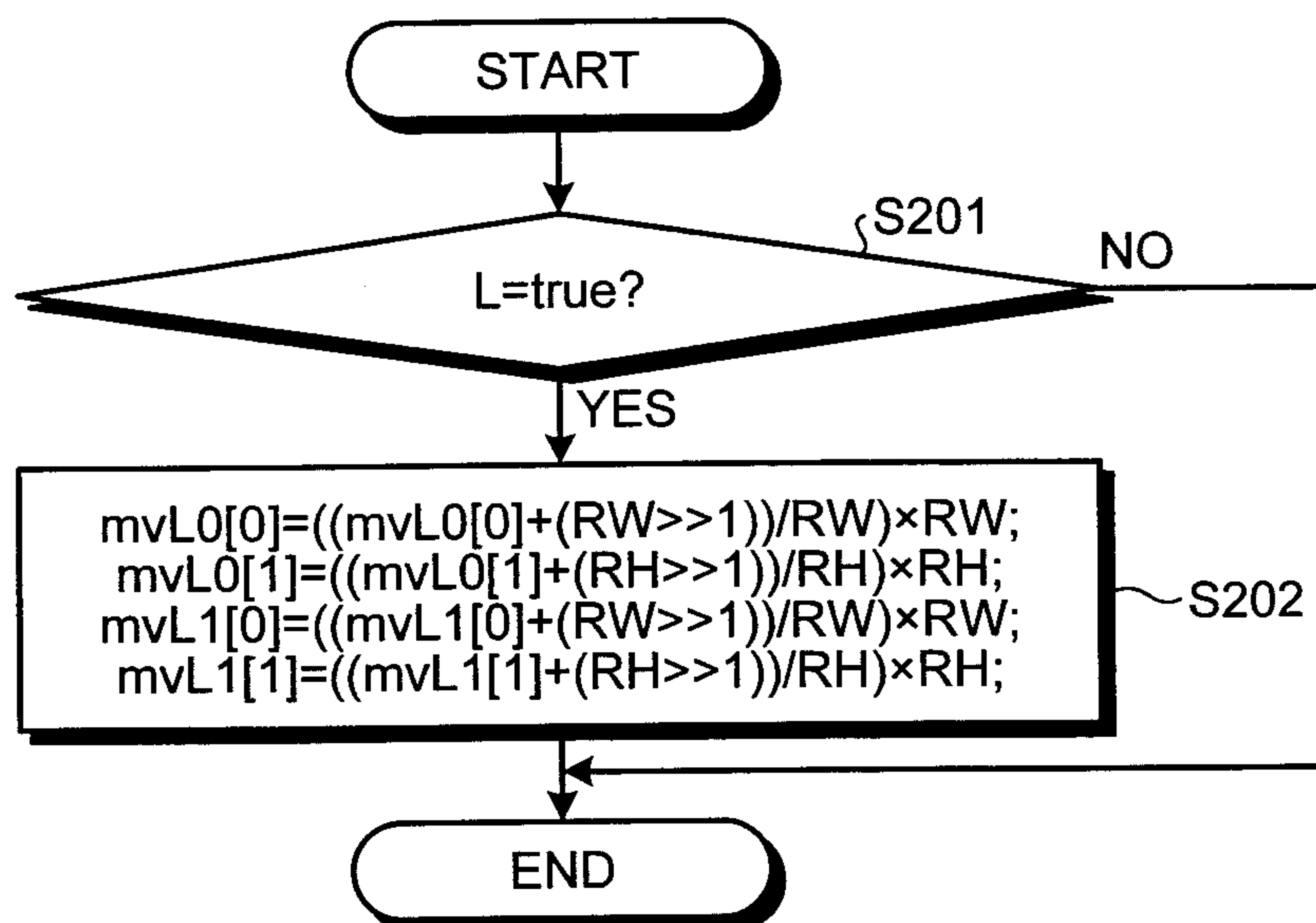
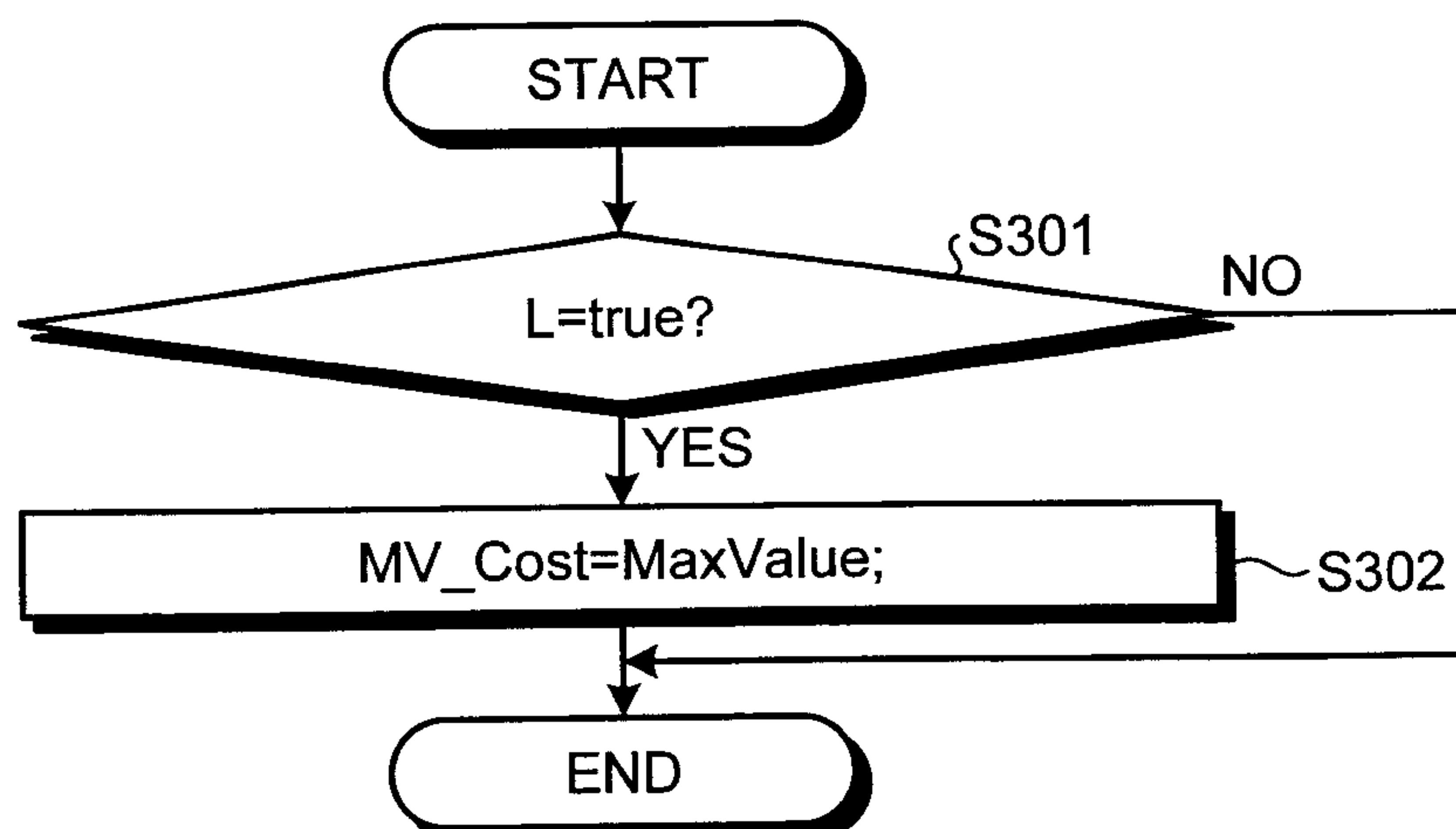


FIG.14



13/13

FIG.15

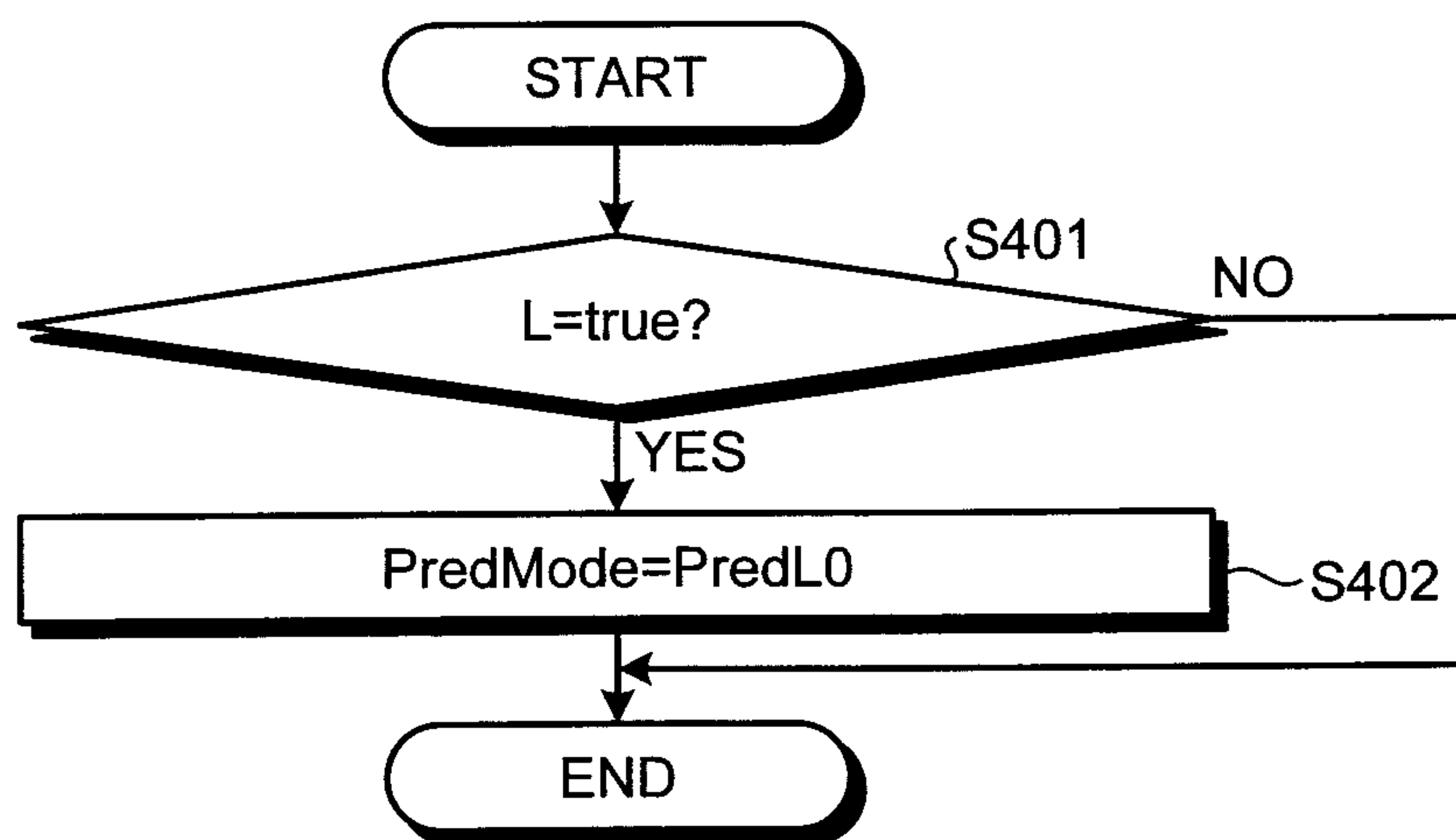


FIG.16

