



(22) Date de dépôt/Filing Date: 1998/07/06

(41) Mise à la disp. pub./Open to Public Insp.: 1999/01/08

(45) Date de délivrance/Issue Date: 2002/02/26

(30) Priorité/Priority: 1997/07/08 (9-196433) JP

(51) Cl.Int.<sup>6</sup>/Int.Cl.<sup>6</sup> G06F 17/30, G11B 23/00

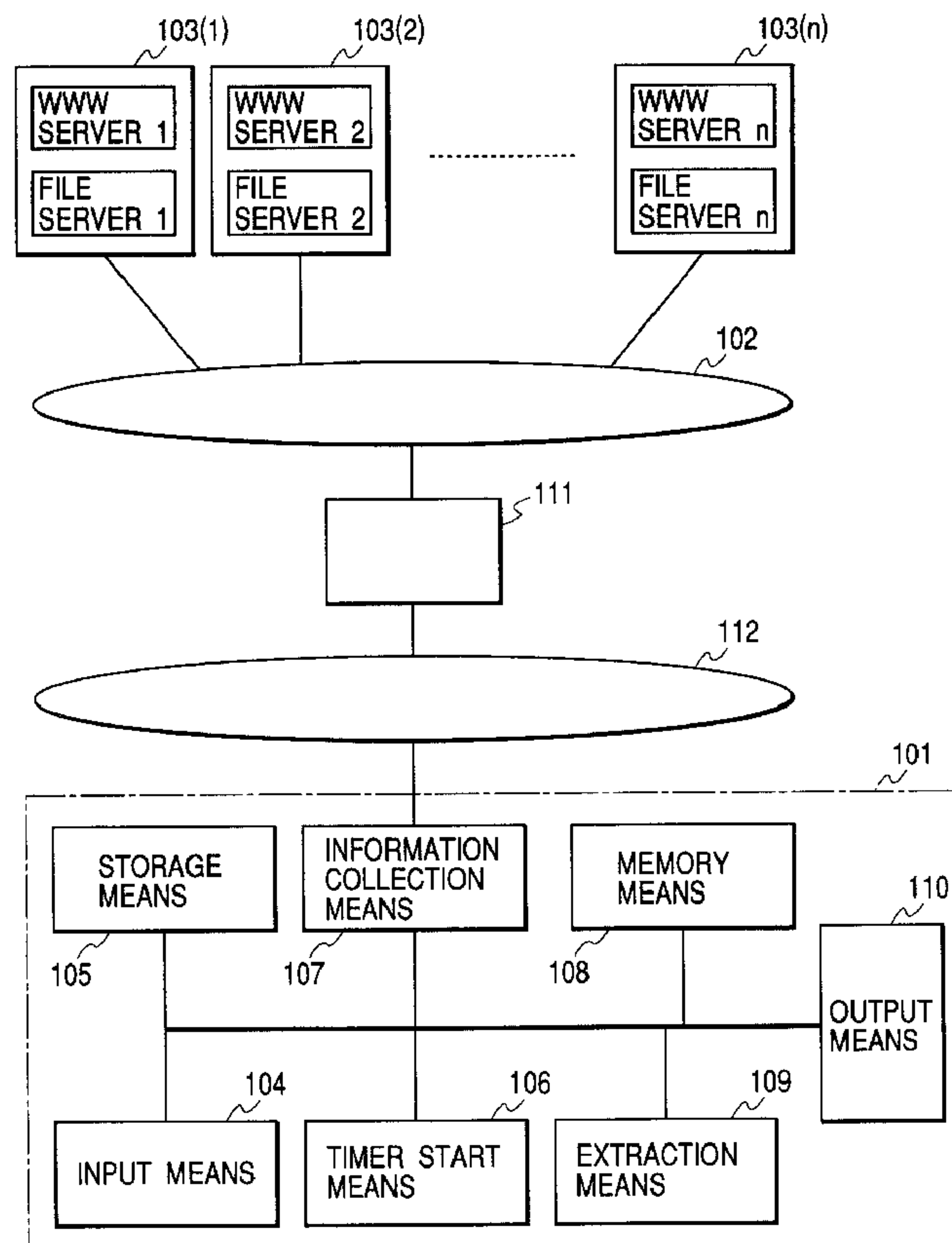
(72) Inventeur/Inventor:  
MATSUI, Hiroshi, JP

(73) Propriétaire/Owner:  
CANON KABUSHIKI KAISHA, JP

(74) Agent: RIDOUT & MAYBEE

(54) Titre : APPAREIL ET METHODE DE RECHERCHE D'INFORMATIONS DANS UN RESEAU

(54) Title: NETWORK INFORMATION SEARCHING APPARATUS AND NETWORK INFORMATION SEARCHING METHOD



(57) Abrégé/Abstract:

A collecting process of data which is managed by each server on a network is first started by information collection means. In the collecting process, data under a point serving as a root is collected while repeating processes such that data at the point



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

serving as a certain root is collected on the basis of access setting information, a link destination is extracted from the collected data, and data on the extracted link destination is further collected. Subsequently, the data which is necessary by the user is extracted by extraction means. The extracted data is outputted to an output destination. Thus, the information which is required by the user can be searched from the information managed and held in a plurality of information processing apparatuses on the network without forcing surplus labor and time to the user.

## ABSTRACT OF THE DISCLOSURE

A collecting process of data which is managed by each server on a network is first started by information collection means. In the collecting process, data under a point serving as a root is collected while repeating processes such that data at the point serving as a certain root is collected on the basis of access setting information, a link destination is extracted from the collected data, and data on the extracted link destination is further collected. Subsequently, the data which is necessary by the user is extracted by extraction means. The extracted data is outputted to an output destination. Thus, the information which is required by the user can be searched from the information managed and held in a plurality of information processing apparatuses on the network without forcing surplus labor and time to the user.

NETWORK INFORMATION SEARCHING APPARATUS  
AND NETWORK INFORMATION SEARCHING METHOD

BACKGROUND OF THE INVENTION

5 Field of the Invention

The invention relates to a network information searching apparatus which is connected to a plurality of information processing apparatuses on a network through the network and is used to search information  
10 required by the user from information managed and held in each of the information processing apparatuses and also relates to a method for such an apparatus.

Related Background Art

In case of searching desired information of the  
15 user from information managed and held by a plurality of computers on a network, it is necessary to repetitively execute an operation to obtain the information and an operation to confirm the contents of the obtained information until necessary information or  
20 information including the necessary information is obtained. That is, it is necessary to repetitively execute an operation for accessing from a computer of the user through the network to a server program operating on the computer connected to the network and  
25 for obtaining information in accordance with a procedure corresponding to the server program after accessing, an operation for reading the contents of the

obtained information and confirming whether the  
necessary information is included in the information or  
not, and an operation for getting next information if  
the necessary information is not included in the  
5 obtained information.

In case of searching necessary information from  
information of a page construction which is managed by  
a WWW (World Wide Web) system constructed by a WWW  
program operating on a computer connected to the  
10 internet and is logically linked by a URL (Uniform  
Resource Locator), the necessary information is  
searched by repeating operations such that the contents  
of an arbitrary page in information of the page  
construction managed are read, if the necessary  
15 information does not exist in the page, characters or a  
figure, namely, what is called an anchor concerned with  
a URL of a page which may include the necessary  
information at a page that is at present opened is  
clicked by a mouse, a corresponding next page is  
20 displayed, and the contents of this page are read.

An example of the operations in case of searching  
necessary information from the information managed by  
the WWW system will now be described with reference to  
Fig. 14. Fig. 14 is a diagram for explaining the  
25 operation example in case of searching the necessary  
information from the information managed by the  
conventional WWW system.



For example, as shown in Fig. 14, the contents of an arbitrary page 8101 in information of a page construction managed by a plurality of WWW servers 81, 82, ..., and 8n are read. If the necessary information  
5 does not exist in this page, characters or a figure, that is, an anchor 8101-1 here concerned with a URL which may include the necessary information at this page is clicked, so that a page 8102 is displayed. The user reads the contents of the page 8102. When he  
10 confirms that the information which is needed by the user does not exist in the page 8102 in a manner similar to the above operation, characters or a figure concerned with the URL of the page which may include the necessary information is clicked by the mouse,  
15 thereby displaying the next page.

However, in case of searching the necessary information from the information managed by the WWW system mentioned above, as characters or a figure concerned with the URL of the page which may include  
20 the necessary information at the page that is at present opened, which characters or figure should be selected often depends on intuition of the user. The user cannot know the relations among the pages. Therefore, it is necessary to repeat the operations  
25 many times and the user is forced to read unnecessary information other than the object for a period of time until the necessary information is searched. Much

surplus labor and surplus working time are required for the operation by the page including the necessary information is found.

5 SUMMARY OF THE INVENTION

It is an object of the invention to provide a network information searching apparatus and a network information searching method which can search information required by the user from information managed and held in a plurality of information processing apparatuses on a network without forcing surplus labor and time to the user.

To accomplish the above object, according to the invention, there is provided a network information searching apparatus which is connected through a network to a plurality of information processing apparatuses on the network and is used to search information required by the user from information stored in each memory means in each of the information processing apparatuses, comprising:

information collection means for collecting the information stored in each of the memory means through the network;

information holding means for holding the information collected by the information collection means;

extraction means for extracting information

corresponding to the information required by the user  
from the information held in the information holding  
means; and

output means for outputting the information  
5 extracted by the extraction means and corresponding to  
the information required by the user.

To accomplish the above object, according to the  
invention, there is provided a network information  
searching method in a network information searching  
10 apparatus which is connected through a network to a  
plurality of information processing apparatuses on the  
network and is used to search information required by  
the user from information stored in each of the memory  
means in each of the information processing  
15 apparatuses, comprising:

an information collecting step of collecting the  
information stored in each of the memory means through  
the network;

an information holding step of holding the  
20 information collected by the information collecting  
step into information holding means;

an extracting step of extracting information  
corresponding to the information required by the user  
from the information held in the information holding  
25 means; and

an output step of outputting the information  
extracted by the extracting step and corresponding to



the information required by the user.

To accomplish the above object, according to the invention, there is provided a storage medium which is connected through a network to a plurality of

5 information processing apparatuses on the network and in which a network information searching program to search information required by the user from information stored in each memory means in each of the information processing apparatuses has been stored,

10 wherein the network information searching program comprises:

an information collecting module for collecting the information stored in each of the memory means through the network;

15 an information holding module for holding the collected information into information holding means;

an extracting module for extracting information corresponding to the information required by the user from the information held in the information holding means; and

20 an output module for outputting the extracted information corresponding to the information required by the user.

To accomplish the above object, according to the invention, there is provided a network information searching system constructed by a plurality of information processing apparatuses and a network

information searching apparatus which are mutually connected through a network,

wherein the network information searching apparatus comprises:

5           information collection means for collecting the information stored in each of the memory means in each of the information processing apparatuses through the network;

          information holding means for holding the  
10   information collected by the information collection means;

          extraction means for extracting information corresponding to the information required by the user from the information held in the information holding  
15   means; and

          output means for outputting the information extracted by the extraction means and corresponding to the information required by the user.

## 20   BRIEF DESCRIPTION OF THE DRAWINGS

          Fig. 1 is a block diagram showing a construction of an embodiment of a network information searching apparatus of the invention;

          Fig. 2 is a diagram showing the contents of  
25   environment setting information which is used in the network information searching apparatus of Fig. 1;

          Fig. 3 is a diagram showing the contents of access

setting information which is used in the network  
information searching apparatus of Fig. 1;

Fig. 4 is a diagram showing the contents of  
extraction condition setting information which is used  
5 in the network information searching apparatus of Fig.  
1;

Fig. 5 is a diagram showing the contents of  
address information which is used in the network  
information searching apparatus of Fig. 1;

10 Fig. 6 is a diagram showing the contents of image  
characteristics information which is used in the  
network information searching apparatus of Fig. 1;

Fig. 7 is a diagram for explaining an example of  
the operation in an information search using the  
15 network information searching apparatus of Fig. 1;

Fig. 8 is a diagram conceptually showing data  
stored in storage means in the network information  
searching apparatus of Fig. 1;

Fig. 9 is a diagram for explaining an example of a  
20 directivity data collection from a WWW system in the  
network information searching apparatus of Fig. 1;

Fig. 10 is a flowchart showing a processing  
procedure in the network information searching  
apparatus of Fig. 1;

25 Fig. 11 is a flowchart showing a processing  
procedure for searching a server program which operates  
on a server computer on a network in the network

information searching apparatus of Fig. 1;

Fig. 12 is a flowchart showing a processing  
procedure in case of collecting information having  
directivity in the network information searching  
5 apparatus of Fig. 1;

Fig. 13 is a flowchart showing a processing  
procedure for extracting image data having designated  
characteristics in the network information searching  
apparatus of Fig. 1; and

10 Fig. 14 is a diagram for explaining an operation  
example in case of searching necessary information from  
information managed by a conventional WWW system.

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

15 An embodiment of the invention will now be  
described hereinbelow with reference to the drawings.

##### (a) First Embodiment:

Fig. 1 is a block diagram showing a construction  
of an embodiment of a network information searching  
20 apparatus of the invention.

As shown in Fig. 1, a plurality of server  
computers 103(m) ( $m = 1, 2, \dots, n$ ) are connected to an  
internet network (hereinafter, referred to as a  
network) 102. An information management system to  
25 manage holding information by server programs such as  
WWW server programs 1 to n, file server programs 1 to  
n, and the like is constructed in each of the server



computers 103(m).

A network information searching apparatus 101 is used as an apparatus for searching information required by the user from the information held in each of the  
5 server computers 103(m). The apparatus 101 is connected to the network 102 through a public line 112 and a provider 111. The network information searching apparatus 101 comprises input means 104, storage means 105, timer start means 106, information collection  
10 means 107, memory means 108, extraction means 109, and output means 110.

The input means 104 inputs various information such as environment setting information, access setting information, extraction condition setting information,  
15 address information, extraction image information, and the like in accordance with the operation by the user. The inputted various information is stored into the storage means 105.

Fig. 2 is a diagram showing the contents of the  
20 environment setting information which is used in the network information searching apparatus of Fig. 1.

As shown in Fig. 2, in the environment setting information, there have been registered: a date, a day of the week, and time of "start"; a date, a day of the  
25 week, and time of "end"; a timer setting to set everyday, every week, or every month in case of periodically starting; a timeout of the connection; the

number of times of retry; a transfer amount upon downloading; and the number of sessions which are simultaneously accessed.

Fig. 3 is a diagram showing the contents of the access setting information which is used in the network information searching apparatus of Fig. 1.

As shown in Fig. 3, in the access setting information, there have been registered: an access setting No.; an address of the server which is used; a classification of the server; a URL serving as a root in case of accessing to the WWW server; a hierarchical layer showing a depth of downloading; a mail address of the user in case of accessing to a mail system; a directory serving as a root in case of accessing to a file server; a user ID; a password; a directivity collection to set whether the directivity collection is performed or not; and a hierarchical layer serving as a unit of discrimination about the directivity collection.

Fig. 4 is a diagram showing the contents of the extraction condition setting information which is used in the network information searching apparatus of Fig. 1.

As shown in Fig. 4, in the extraction condition setting information, there have been registered: an extraction condition setting No.; an access setting No. concerned with the access setting No. of the access

setting information; a search extraction to set whether  
the extraction by the search is performed or not; a  
synonym extraction to set whether the extraction by a  
synonym is performed or not; an analogue extraction to  
5 set whether the extraction by an analogue is performed  
or not; a search object to designate a search object;  
search conditions 1 to n to set a search keyword and  
the like; a data type extraction to set whether the  
extraction by a data type is performed or not; data  
10 types 1 to n to designate a type of data to be  
extracted; an image extraction to set whether the  
extraction by characteristics of the image data is  
performed or not; image conditions 1 to n to designate  
characteristics of the image data to be extracted; an  
15 output destination designation to set an output  
destination; an address No. in case of outputting to a  
mail or FAX; and extra condition data to set whether  
the data which is not congruous with the extraction  
conditions is left or not.

20 Fig. 5 is a diagram showing the contents of the  
address information which is used in the network  
information searching apparatus of Fig. 1. Fig. 6 is a  
diagram showing the contents of the image  
characteristics information.

25 As shown in Fig. 5, a mail or an FAX telephone  
number of the address has been registered in the  
address information. As shown in Fig. 6, an image



condition No. and image characteristics information have been registered in the image extraction information.

The information collection means 107 comprises means for accessing to the corresponding server computer (m) on the basis of the access setting information and collecting information held by each server computer (m) through the network 102. The information collected by the information collection means 107 is held in the memory means 108. The extraction means 109 extracts the information corresponding to the information required by the user on the basis of the extraction condition setting information set by the user from the information held in the memory means 108. The extracted information is outputted to the output destination designated by the output destination designation included in the extraction condition setting information by the output means 110. The timer start means 106 comprises means for starting the information collection means 107 on the basis of the information included in the environment setting information.

Subsequently, a series of operations in case of searching the information which is needed by the user will now be described by using the network information searching apparatus 101 with reference to Figs. 1 to 5 and 7. Fig. 7 is a diagram for explaining an example



of the operation in the information search using the network information searching apparatus of Fig. 1. The series of operations will now be described with respect to an example in the case where the apparatus is  
5 connected to the network 102 at night when the network 102 is free, a series of WWW pages designated by the user is collected, the pages which are needed by the user are extracted from the collected pages, and the extracted pages are browsed in the daytime.

10 It is assumed that the information which is needed by the user, for example, the pages including the information regarding a printing device are a page 718 in a server 710 of A newspaper company and a page 7n4 in a server 7nn of N newspaper company and that a page  
15 of an index having a logical link to each newspaper company is a page 702 under the management of the WWW server 701 of CC Co. Ltd.

In this case, first, the user sets various information through the input means 104. In the access  
20 setting information, as shown in the access setting No. 1 in Fig. 3, a URL of the page 702 as a URL serving as a root is set to 3 as a hierarchical layer which traces the link. In case of storing the extracted information into a homepage and confirming, in the extraction  
25 condition setting information, as shown in the extraction condition No. 1 in Fig. 4, 1 is set into the access setting No. and the storage device (storage

means 105 here) is set to the output destination,  
respectively. Further, for example, in case of setting  
so as to start the information collection means 1 at 3  
a.m. when it is presumed that the network 102 and WWW  
5 server 701 will be free, as shown in Fig. 2, the start  
time in the environment information setting information  
is set to 3 a.m..

When the time of 3 a.m. comes, the information  
collection means 107 is started by the timer start  
10 means 106. The information collection means 107 refers  
to the access setting information held in the storage  
means 105, connects to the provider 111 through the  
public line 112, and connects to the network 102 via  
the provider 111. After it was connected to the  
15 network 102, first, the data at the page 702 under the  
management of the WWW server 701 of CC Co., Ltd. is  
collected. The information of the linked URL is  
extracted from the collected data. The data at pages  
711 to 7n1 on the linked destination side is collected  
20 from the page 702 on the basis of the extracted link  
information. The information of the linked URL is  
extracted from those pages. By repeating the processes  
for collecting the page data of further lower  
hierarchical layers on the basis of the extracted link  
25 information, the data as much as four hierarchical  
layers is collected. After completion of the  
collection of the data, the connection to the provider

111 through the public line 112 is disconnected. The data collected by the information collection means 107 is held in the memory means 108.

5 The data corresponding to the information which is needed by the user is extracted by the extraction means 109. In the extracting process, with reference to the extraction condition setting information, the data held in the memory means 108 is discriminated to see whether a keyword of "printing device", a synonym, or an  
10 analogue is included in the whole text or not on a page unit basis. If there is a page including the keyword, synonym, or analogue, the data of this page is stored into the storage means 105.

After the data was stored, the user displays the  
15 data stored in the storage means 105 to display means (not shown) and confirms the contents of the data.

The data stored in the storage means 105 will now be described with reference to Fig. 8. Fig. 8 is a diagram conceptually showing the data stored in the  
20 storage means of the network information searching apparatus of Fig. 1.

As shown in Fig. 8, page data 9101, 9102, and 9103 extracted by the extracting process is stored in the storage means 105. The page data 9101 is the data of  
25 the page having index information of the extraction result. The page data 9102 is the data of the page corresponding to the page 718 shown in Fig. 7. The



page data 9103 is the data of the page corresponding to the page 7n4 shown in Fig. 7. Therefore, it is sufficient for the user to confirm the page data 9102 and 9103. It will be understood that the work can be remarkably reduced as compared with a work to confirm the pages of the lower hierarchical layers in which the page 702 in Fig. 7 is the top hierarchy.

A processing procedure in the network information searching apparatus 101 will now be described with reference to Fig. 10. Fig. 10 is a flowchart showing a processing procedure in the network information searching apparatus of Fig. 1.

First, the processing routine is started in step S1. The "start" of the processing routine denotes that the information collecting process by the information collection means 107 is started. As a starting method, there is a start by the timer start means 106 or a start by the operation of the input means 104.

In step S2, the collecting process of the network information by the information collection means 107 is subsequently executed. In the collecting process, the data under the point serving as a root is collected up to the set hierarchical layer while repeating processes for collecting the data at a point serving as a certain root on the basis of the access setting information, extracting the link destination from the collected data, and further collecting the data on the extracted



link destination. The "point serving as a root" here denotes a page shown by the URL in the WWW system, a directory of the file system in the file server system on the network, or a mail address in the mail system.

5           In step S3, the data which is needed by the user is extracted by the extraction means 109. The extraction condition which is used at the time of data extraction is held as extraction condition setting information in the storage means 105 as mentioned  
10       above. As an extracting method of the data, there is an extraction of data including the image set by the user, an extraction of data of the type set by the user, an extraction of data including the data set by the user, an extraction by the keyword set by the user,  
15       an extraction by the synonym of the keyword set by the user, or an extraction by the analogue of the keyword set by the user.

          The extraction of data of the type set by the user denotes that, for example, in case of the data of the  
20       page under the management of the WWW system, only the data of the designated type is extracted from the page. As a type of data, there is text data, image data, audio data, or binary data.

          The extraction of data including the data set by  
25       the user denotes that, for example, in case of the data of the page under the management of the WWW system, all of the data of the page including the data of the

designated type is extracted from the page, and that in case of the file, the file itself including the data of the designated type is extracted.

5       The extraction by the keyword, synonym of the keyword, or analogue of the keyword denotes that on the basis of the keyword and its synonym or analogue, the search is performed for the title of the data, file name, URL name, and the whole sentences of the text as targets, and the data including the keyword and its  
10       synonym or analogue is extracted. Whether the extracted data is congruous with the search condition or not is determined by designating by the user after the data was extracted.

15       In step S4, the extracted data is outputted. In the data output, the extracted data is outputted to the output destination designated by the output destination designation included in the extraction condition setting information. As an output destination, it is possible to set the storage means 105, printing device,  
20       display means, arbitrary FAX number, arbitrary mail address, or the like.

25       In step S5, the confirmation of the user for the outputted information is performed. The processing routine is finished.

#### (b) Second Embodiment

In step S2 mentioned above, the process has been

performed on the assumption of the case where the information such as server address, URL, or the like in which the point serving as a root exists has been set by the access setting information, namely, the case  
5 where the user knows the information of the server address, URL, or the like as a prerequisite. However, processes in the case where the user does not know such information will now be described with reference to Fig. 11. Fig. 11 is a flowchart showing a procedure  
10 for processes for searching a server program operating on the server computer on the network in the network information searching apparatus of Fig. 1.

First in step S21, a response request is issued to each of the server computers 103(m) on the network 102.  
15 In step S22, a check is made to see if there is a response to the response request from each server computer 103(m). If there is a response, step S23 follows and a request is issued to the server on the server computer 103(m) which responded. In step S24, a  
20 check is made to see if there is a response to the request from the corresponding server. If there is a response to the request from the corresponding server, the processing routine advances to a process to search the data which is needed by the user while using the  
25 server as a root. If the target data does not exist in the hierarchical data constructed under the dominant of the root, the processing routine is returned to step



S21 and the process to search the next server is performed.

The order of the server computers 103(m) to generate the response request in step S21 is determined by preferentially using the method selected by the user among the well-known method of preferentially using the typical address, a method of preferentially using the address near the address of the user, and a method of accessing at random.

10

(c) Third Embodiment:

In step S2, a processing procedure in case of collecting the information having directivity will now be described with reference to Fig. 12. Fig. 12 is a flowchart showing the processing procedure in case of collecting the information having directivity in the network information searching apparatus of Fig. 1.

15

First in step S31, the data as much as the hierarchical layers set in the access setting information is collected. In step S32, a grade congruous with the extraction condition is numerically evaluated.

20

In step S33, the hierarchical data of data located under the page having a high grade congruous with the extraction condition is collected by only an amount of the directivity collection hierarchical layers set in the access setting information. In step S34, a check

25



- 22 -

is made to see if an end condition has been satisfied.  
If the end condition is not satisfied, the processing  
routine is returned to step S32. Processes from step  
S32 are repetitively executed by the end condition is  
5 satisfied. When the end condition is satisfied, the  
processing routine is finished.

The collection of the information having the  
directivity will now be described with reference to  
Fig. 9 with respect to an example in the case of  
10 collecting the data from the WWW system. Fig. 9 is a  
diagram for explaining an example of the directivity  
data collection from the WWW system in the network  
information searching apparatus of Fig. 1.

In Fig. 9, in the case where the hierarchical  
15 layer to collect the data in a lump is labelled as "1"  
while using a page 10101 as a root, first, the data of  
pages 10102 to 10104 of the hierarchical layers under  
the page 10101 is collected, a grade congruous with the  
extraction condition is numerically evaluated, and it  
20 is assumed that the page in which the congruous grade  
is the highest is the page 10103. In this instance,  
the data of pages 10106 to 10108 under the page 10103  
is collected. Similar processes are repeated. When it  
is assumed that the pages in which the grade congruous  
25 with the extraction condition is high are the page  
10106 among the pages 10106 to 10108 and a page 10202  
among pages 10202 to 10204, the pages included in the

range of a page 11111 to be collected are collected.

(d) Fourth Embodiment:

In step S3, the case of extracting the image data  
5 having the designated characteristics will now be  
described with reference to Fig. 13. Fig. 13 is a  
flowchart showing a processing procedure for extracting  
the image data having designated characteristics in the  
network information searching apparatus of Fig. 1.

10 First in step S41, the image data is extracted  
from the information collected in step S2 (shown in  
Fig. 10). In step S42, characteristics of the  
extracted image data are extracted. In step S43, by  
comparing the extracted characteristics of the image  
15 with the image extraction information (shown in Fig. 6)  
set by the user, the image data having the  
characteristics of the image data which is needed by  
the user is extracted.

As characteristics of the image which is set by  
20 the user, there are an outline and a color such as  
human object, house, or the like solely having certain  
characteristics in case of an outline, scenery  
characteristics such as mountain, plain, sea, four  
seasons, or the like in case of a scene, and the like.

25 As mentioned above, in the network information  
searching apparatus 101, the data held by the server  
system of each server computer 103(m) is collected

through the network 102. The data corresponding to the data required by the user is extracted from the collected data. The extracted data is outputted to the set output destination. Therefore, the data which is  
5 needed by the user can be searched from the data managed and held by the server system of each of the server computers 103(m) on the network 102 without forcing surplus labor and time to the user.

Data can be automatically collected in a time zone  
10 when the network is free, the communication time can be reduced, and the communication costs can be reduced.

Further, since the extracted data is outputted to the designated output destination, for example, it is possible to easily set a mode to print and output the  
15 pages which are seen everyday by a predetermined time in the morning, a mode to output them as a mail to himself, a mode to transmit them to an FAX apparatus of a predetermined address, or the like. The extracted information can be confirmed in a desired output form  
20 of the user.

In the embodiment, although the example in which the network information searching apparatus 101 is constructed as a single apparatus has been shown, the network information searching apparatus 101 can be also  
25 constructed by a personal computer. In this case, there is used a personal computer comprising at least: a keyboard or a mouse constructing the input means 104;



a display; an external memory device such as hard disk or the like constructing the storage means 105 or memory means 108; and an interface to connect to the provider 111 through the public line. A network  
5 information searching program including an information collecting module constructing the information collection means 107, an information holding module for holding the collected data into the external memory device, an extracting module constructing the  
10 extraction means 109, and an output module constructing the output means 110 is stored in the external memory device. By reading out and executing the network information searching program as necessary, the network information searching apparatus is constructed. The  
15 network information searching program is supplied by a memory medium such as FD (floppy disk), CD-ROM, or the like.

As described above, according to the first to fourth embodiments, the information which is required  
20 by the user can be searched from the information managed and held by the information processing program of a plurality of information processing apparatuses on the network without forcing surplus labor and time to the user.

25 The data can be automatically collected in a time zone when the network is free, the communication time can be reduced, and the communication costs can be

decreased.

It is also possible to easily connect to the information processing apparatus in which the connection is permitted from a plurality of information processing apparatuses on the network.

The information required by the user can be properly extracted.

It is possible to easily set the mode to print and output the pages which are seen everyday by a predetermined time, the mode to output them as a mail to himself, the mode to transmit them to the FAX apparatus of a predetermined address, or the like. The extracted information can be confirmed in a desired output form of the user can be confirmed.

- 27 -

## WHAT IS CLAIMED IS:

1. A network information searching apparatus which is connected through a network to a plurality of information processing apparatuses on the network and  
5 is used to search information required by a user from information stored in each memory means in each of said information processing apparatuses, comprising:

information collection means for collecting the information stored in each of said memory means through  
10 said network;

information holding means for holding the information collected by said information collection means;

extraction means for extracting information  
15 corresponding to the information required by said user from the information held in said information holding means; and

output means for outputting the information extracted by said extraction means and corresponding to  
20 the information required by said user.

2. An apparatus according to claim 1, further comprising timer start means for starting said information collection means at a start time set by  
25 said user.

3. An apparatus according to claim 1, wherein said



information collection means has detection means for  
detecting an address of the information processing  
apparatus in which a connection is permitted for said  
user among said information processing apparatuses and  
5 an information management system of said address.

4. An apparatus according to claim 1, wherein each  
of said memory means is controlled by an information  
management program such as World Wide Web server  
10 program, file server program for a file server system,  
or the like.

5. An apparatus according to claim 1, wherein said  
information collection means collects the information  
15 stored in each of said memory means with directivity  
for preferentially tracing information concerned with  
the information required by said user at a  
predetermined probability.

20 6. An apparatus according to claim 1, wherein when  
the information required by said user is image data,  
said extraction means extracts image data from the  
information held in said information holding means,  
extracts characteristics of said extracted image data,  
25 compares said extracted image data with characteristics  
of the image data required by said user, and extracts  
image data corresponding to the image data required by

said user in accordance with a result of said comparison.

7. An apparatus according to claim 1, wherein said  
5 extraction means analyzes a data type of the  
information held in said information holding means,  
compares a result of said analysis with a data type of  
the information required by said user, and extracts  
information corresponding to the information required  
10 by said user in accordance with a result of said  
comparison.

8. An apparatus according to claim 1, wherein said  
extraction means analyzes a data type of the  
15 information held in said information holding means,  
compares a result of said analysis with a data type of  
the information required by said user, and extracts  
information including information of a data type  
corresponding to the data type of the information  
20 required by said user in accordance with a result of  
said comparison.

9. An apparatus according to claim 1, wherein said  
extraction means analyzes an attribute of the  
25 information held in said information holding means,  
compares a result of said analysis with an attribute of  
the information required by said user, and extracts

- 30 -

information corresponding to the information required by said user in accordance with a result of said comparison.

5        10. An apparatus according to claim 1, wherein said extraction means performs a search to the information held in said information holding means on the basis of a keyword included in a text portion of said information and a synonym or an analogue of said  
10       keyword and extracts information corresponding to the information required by said user by said search.

11. An apparatus according to claim 1, wherein said extraction means extracts information corresponding to  
15       the information required by said user and forms an index of information concerned in an extracting step of said extracted information, and said output means outputs said index formed together with the information extracted by said extraction means.

20       12. An apparatus according to claim 1, wherein said output means outputs said extracted information to an apparatus which can output said extracted information in a desired output form of said user in accordance  
25       with a selection of said user.

13. A network information searching method of a



- 31 -

network information searching apparatus which is  
connected through a network to a plurality of  
information processing apparatuses on the network and  
is used to search information required by a user from  
5 information stored in each memory means in each of said  
information processing apparatuses, comprising:

an information collecting step of collecting the  
information stored in each of said memory means through  
said network;

10 an information holding step of holding said  
collected information into information holding means;

an extracting step of extracting information  
corresponding to the information required by said user  
from the information held in said information holding  
15 means; and

an output step of outputting said extracted  
information corresponding to the information required  
by said user.

20 14. A method according to claim 13, further  
comprising a step of instructing so as to start the  
collection of the information stored in each of said  
memory means at a start time set by said user.

25 15. A method according to claim 13, wherein each of  
said memory means is controlled by an information  
management program such as World Wide Web server

- 32 -

program, file server program for a file server system,  
or the like.

16. A method according to claim 13, wherein in said  
5 information collecting step, the information stored in  
each of said memory means is collected with directivity  
for preferentially tracing information concerned with  
the information required by said user at a  
predetermined probability.

10

17. A method according to claim 13, wherein in said  
extracting step, when the information required by said  
user is image data, image data is extracted from the  
information held in said information holding means,  
15 characteristics of said extracted image data are  
extracted, said extracted image data is compared with  
characteristics of the image data required by said  
user, and image data corresponding to the image data  
required by said user is extracted in accordance with a  
20 result of said comparison.

18. A method according to claim 13, wherein in said  
extracting step, a data type of the information held in  
said information holding means is analyzed, a result of  
25 said analysis is compared with a data type of the  
information required by said user, and information  
corresponding to the information required by said user

is extracted in accordance with a result of said comparison.

19. A method according to claim 13, wherein in said  
5 extracting step, a data type of the information held in  
said information holding means is analyzed, a result of  
said analysis is compared with a data type of the  
information required by said user, and information  
including information of a data type corresponding to  
10 the data type of the information required by said user  
is extracted in accordance with a result of said  
comparison.

20. A method according to claim 13, wherein in said  
15 extracting step, an attribute of the information held  
in said information holding means is analyzed, a result  
of said analysis is compared with an attribute of the  
information required by said user, and information  
corresponding to the information required by said user  
20 is extracted in accordance with a result of said  
comparison.

21. A method according to claim 13, wherein in said  
extracting step, a search is executed to the  
25 information held in said information holding means on  
the basis of a keyword included in a text portion of  
said information and a synonym or an analogue of said



- 34 -

keyword and information corresponding to the information required by said user by said search is extracted.

22. A method according to claim 13, wherein in said extracting step, information corresponding to the information required by said user is extracted and an index of information concerned in an extracting step of said extracted information is formed, and in said output step, said formed index is outputted together with the information extracted by said extracting step.

23. A method according to claim 13, wherein in said output step, said extracted information is outputted to an apparatus which can output said extracted information in a desired output form of said user in accordance with a selection of said user.

24. A storage medium which is connected through a network to a plurality of information processing apparatuses on the network and in which a network information searching program to search information required by a user from information stored in each of memory means arranged in each of said information processing apparatus has been stored,

wherein said network information searching program comprises computer-readable program codes of:

executing an information collecting of collecting the information

- 35 -

stored in each of said memory means through said network;

executing an information holding of holding said collected information into information holding means;

executing an extracting of extracting information corresponding to the information required by said user from the information held in said information holding means; and

executing an output of outputting said extracted information corresponding to the information required by said user.

25. A network information searching system constructed by a plurality of information processing apparatuses and a network information searching apparatus which are mutually connected through a network, wherein said network information searching apparatus comprises:  
information collection means for collecting information stored in each memory means in each of said information processing apparatuses through said network;

information holding means for holding the information collected by said information collection means;

- 36 -

extraction means for extracting information  
corresponding to the information required by said user  
from the information held in said information holding  
means; and

5        output means for outputting the information  
extracted by said extraction means and corresponding to  
the information required by said user.

Ridout & Maybee  
Suite 2400  
One Queen Street East  
Toronto, Canada M5C 3B1  
Patent Agents of the Applicant



1 / 13

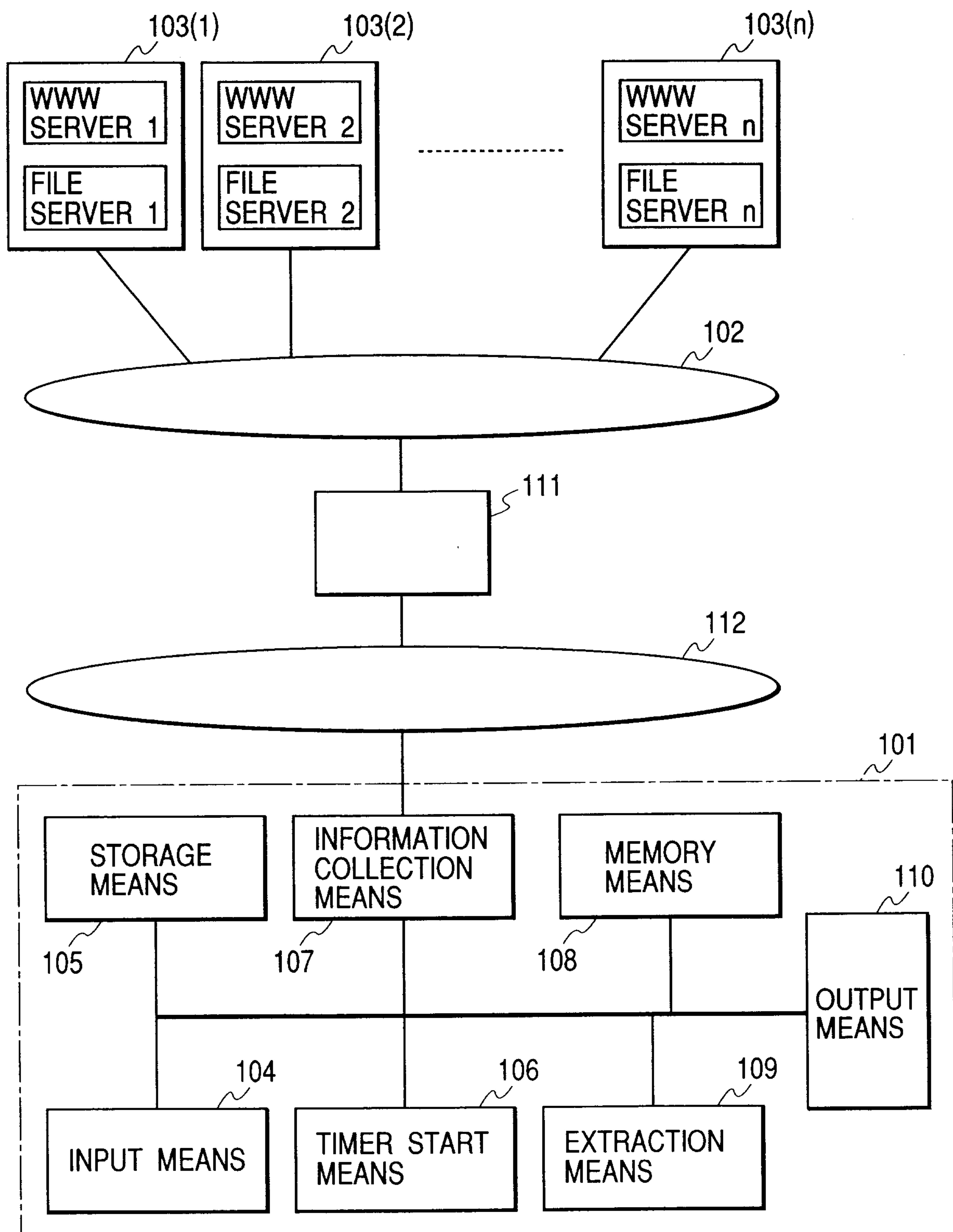
*FIG. 1*



FIG. 3

ACCESS SETTING	SETTING 1	SETTING 2	SETTING 3	SETTING 4	
SERVER ADDRESS				mail.or.jp	
SERVER CLASSIFICATION	WWW	WWW	WWW	MAIL	
URL	URL OF PAGE 702	www.b.co.jp	www.8101.co.jp		
HIERARCHICAL LAYER	3	3	3		
MAIL ADDRESS				aa@mail.or.jp	
DIRECTORY					
USER ID					
PASSWORD					
DIRECTIVITY COLLECTION	NO	NO	NO	NO	
DIRECTIVITY COLLECTION LAYER					



FIG. 4

EXTRACTION CONDITION No.	1	2	3	4	5	6
ACCESS SETTING No.	1	2	1	1	3	4
SEARCH EXTRACTION	YES	NO	NO	NO	YES	NO
SYNONYM EXTRACTION	NO	NO	NO	NO	YES	NO
ANALOGUE EXTRACTION	NO	NO	NO	NO	YES	NO
SEARCH OBJECT	ALL	—	—	—	ALL	—
SEARCH CONDITION 1	PRINTING DEVICE	—	—	—	CANON	—
	—	—	—	—	—	—
SEARCH CONDITION n	—	—	—	—	—	—
DATA TYPE EXTRACTION	NO	NO	NO	NO	NO	NO
DATA TYPE 1	—	—	—	—	—	—
	—	—	—	—	—	—
DATA TYPE n	—	—	—	—	—	—
IMAGE EXTRACTION	NO	NO	NO	NO	NO	NO
IMAGE CONDITION 1	—	—	—	—	—	—
	—	—	—	—	—	—
IMAGE CONDITION n	—	—	—	—	—	—
OUTPUT DESTINATION DESIGNATION	STORAGE DEVICE	PRINTING DEVICE	MAIL	FAX	STORAGE DEVICE	PRINTING DEVICE
ADDRESS No.			1	2		
EXTRA CONDITION DATA			1	2		

5 / 13

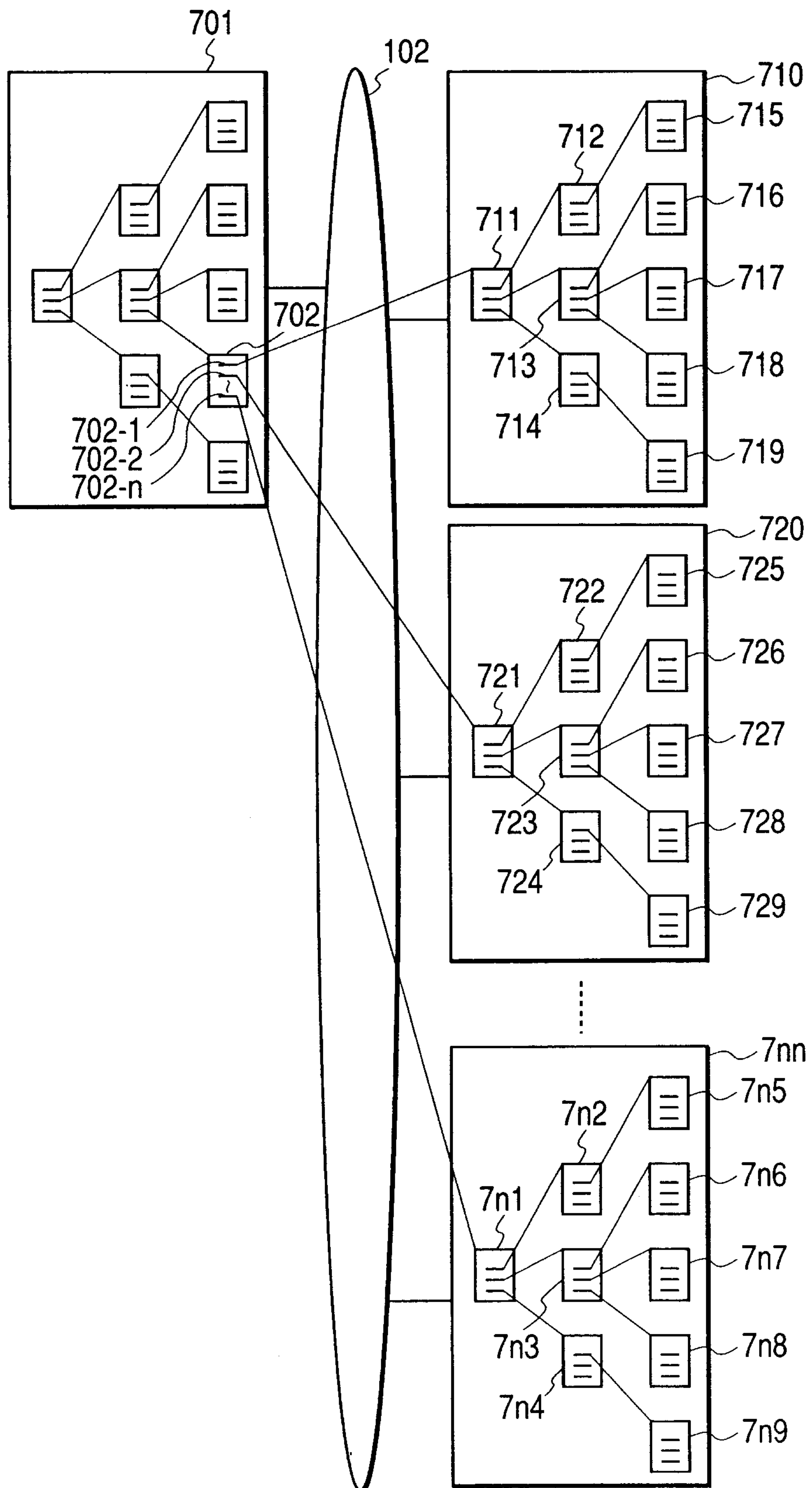
**FIG. 5**

ADDRESS No.	MAIL ADDRESS	FAX No.	
1	aaa@bbb.or.jp		
2		03-1234-1234	
3			

**FIG. 6**

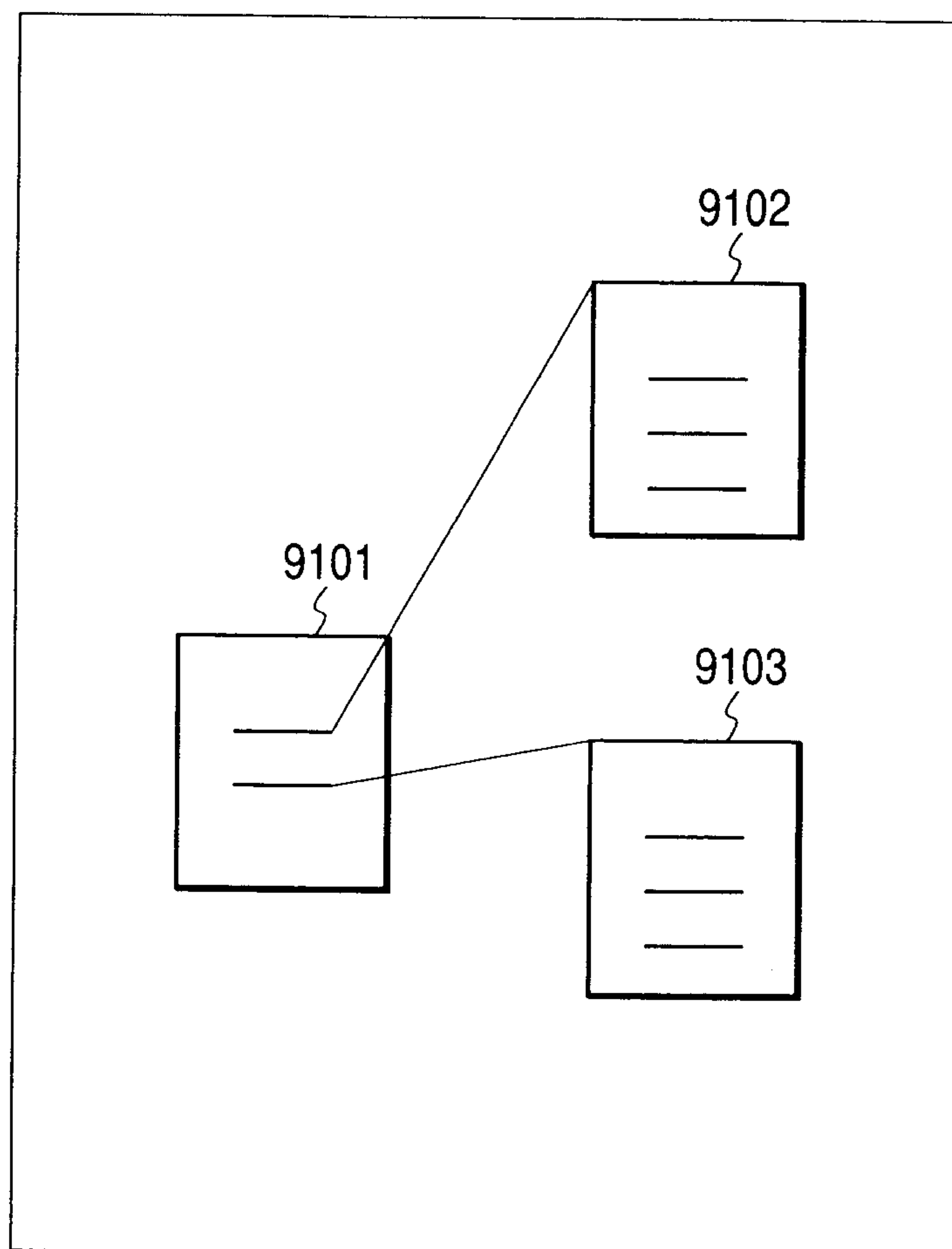
[illegible]

6 / 13

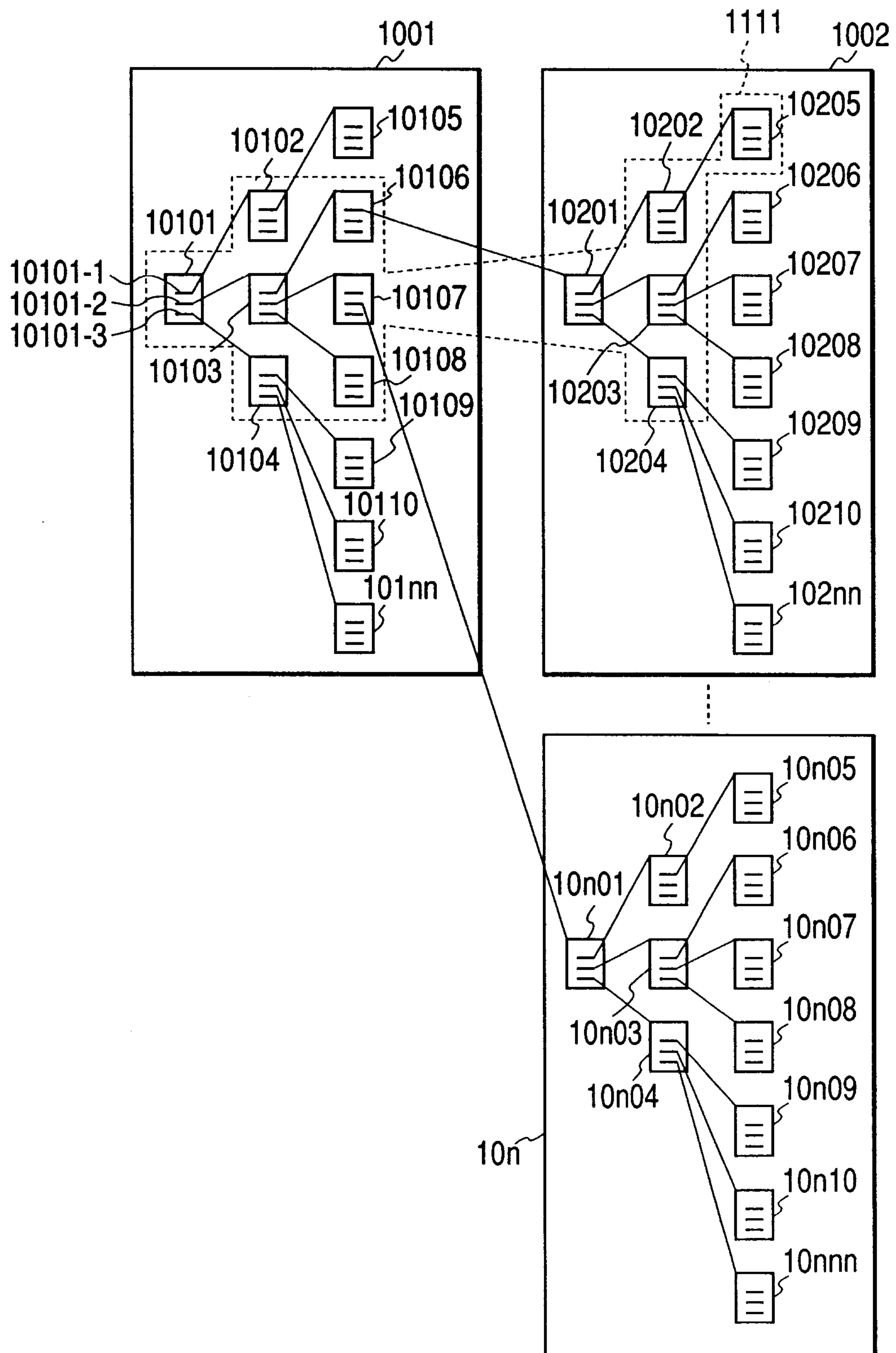
**FIG. 7**

7/13

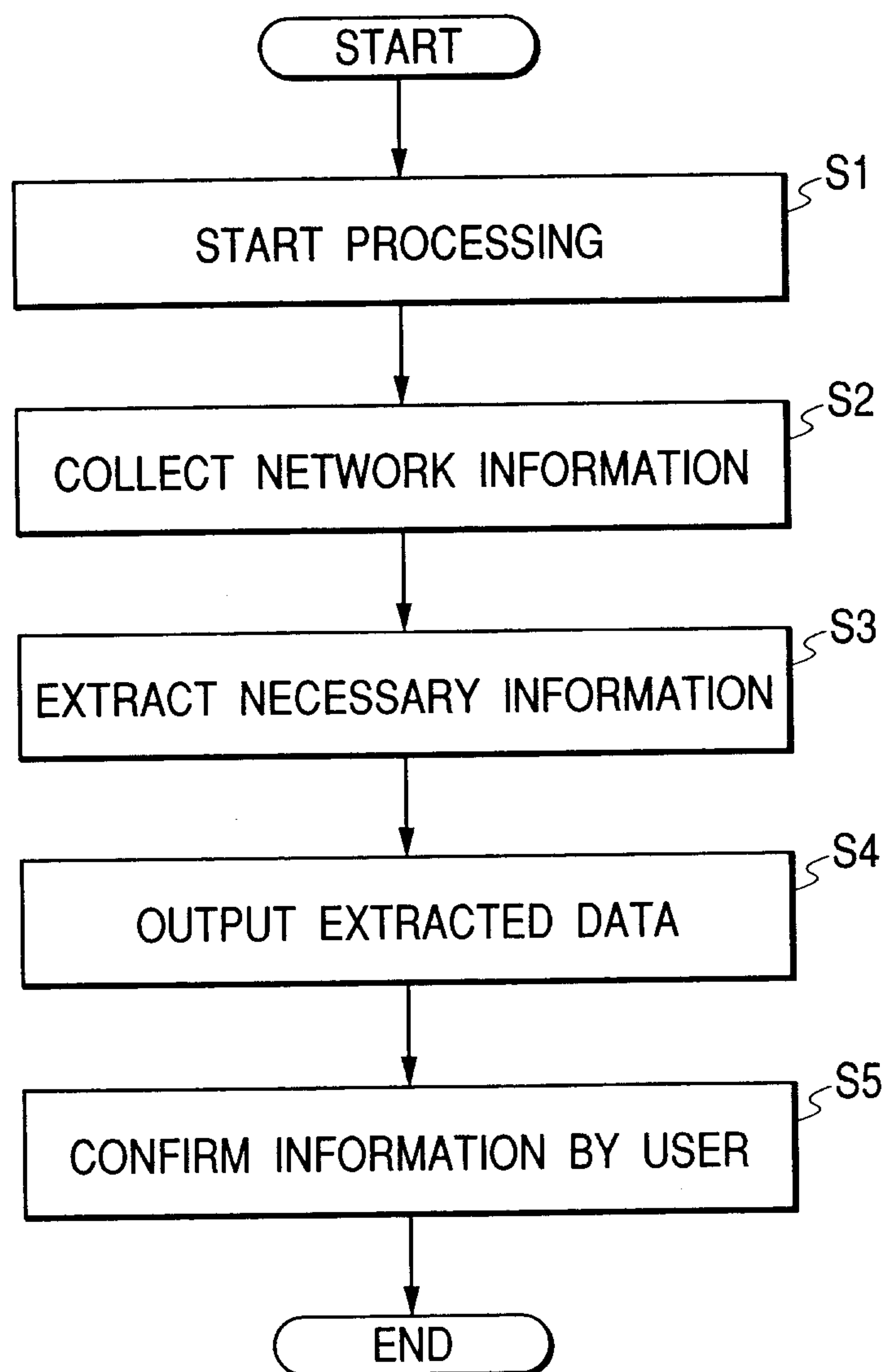
*FIG. 8*



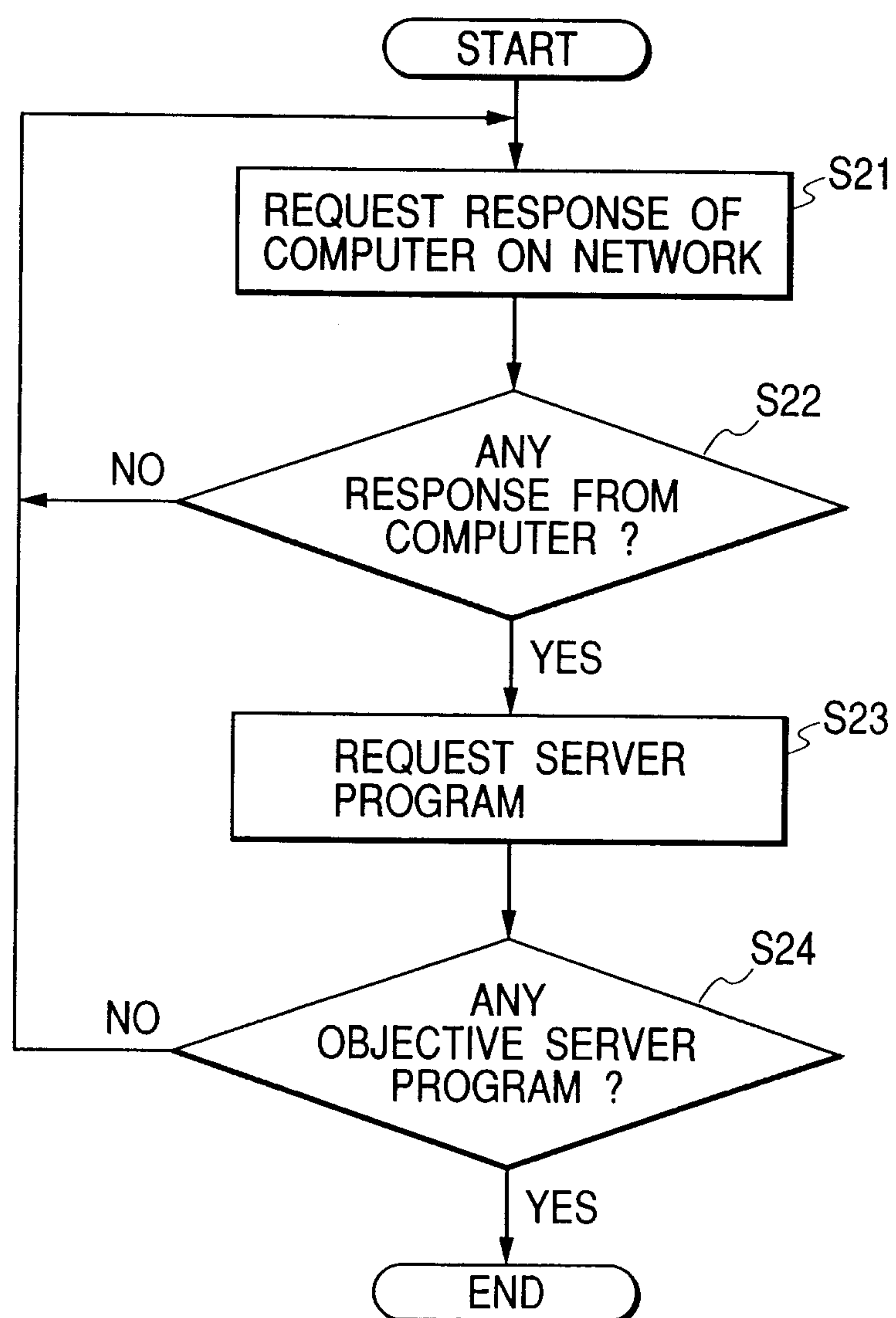


**FIG. 9**

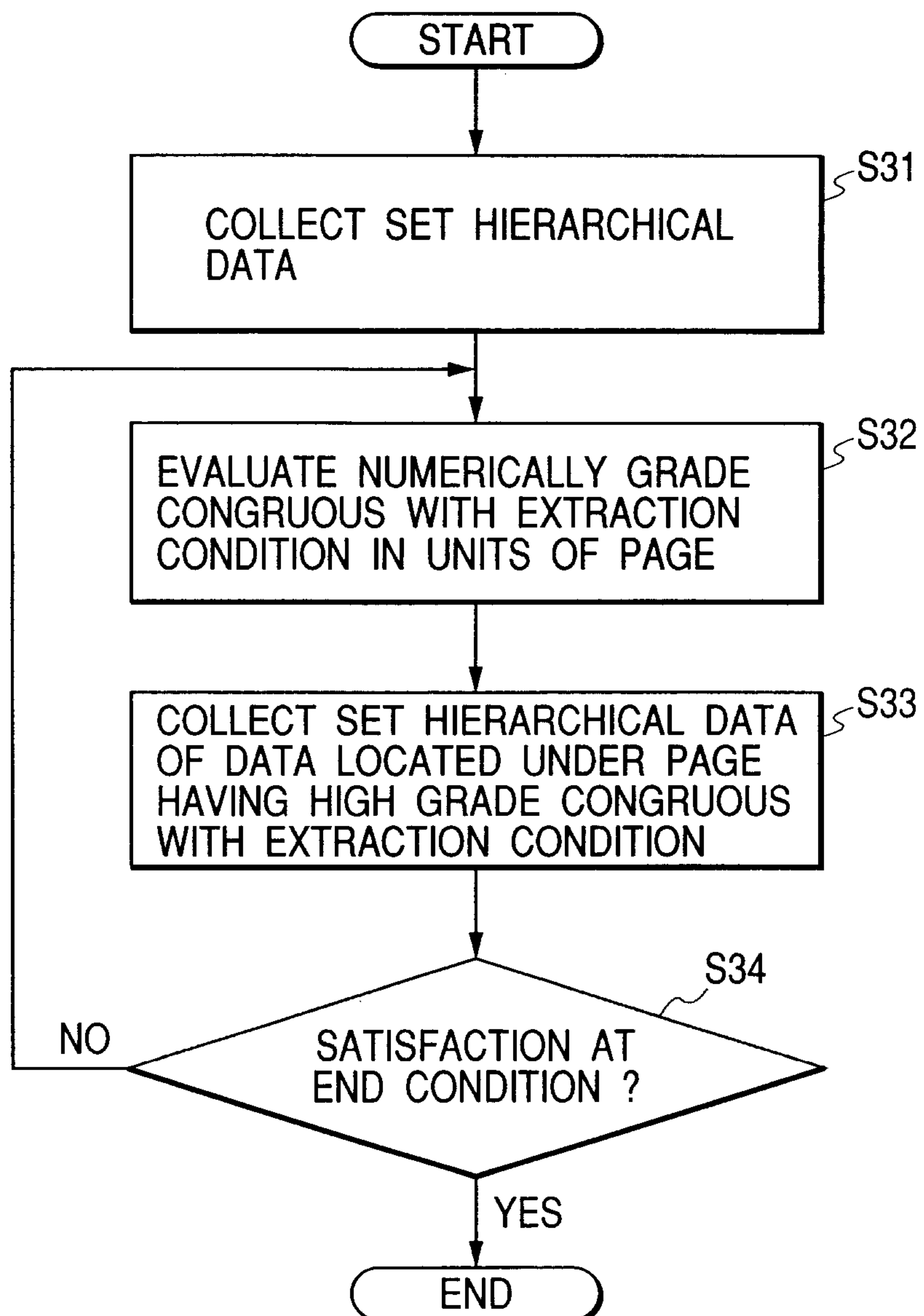
9/13

**FIG. 10**

10 / 13

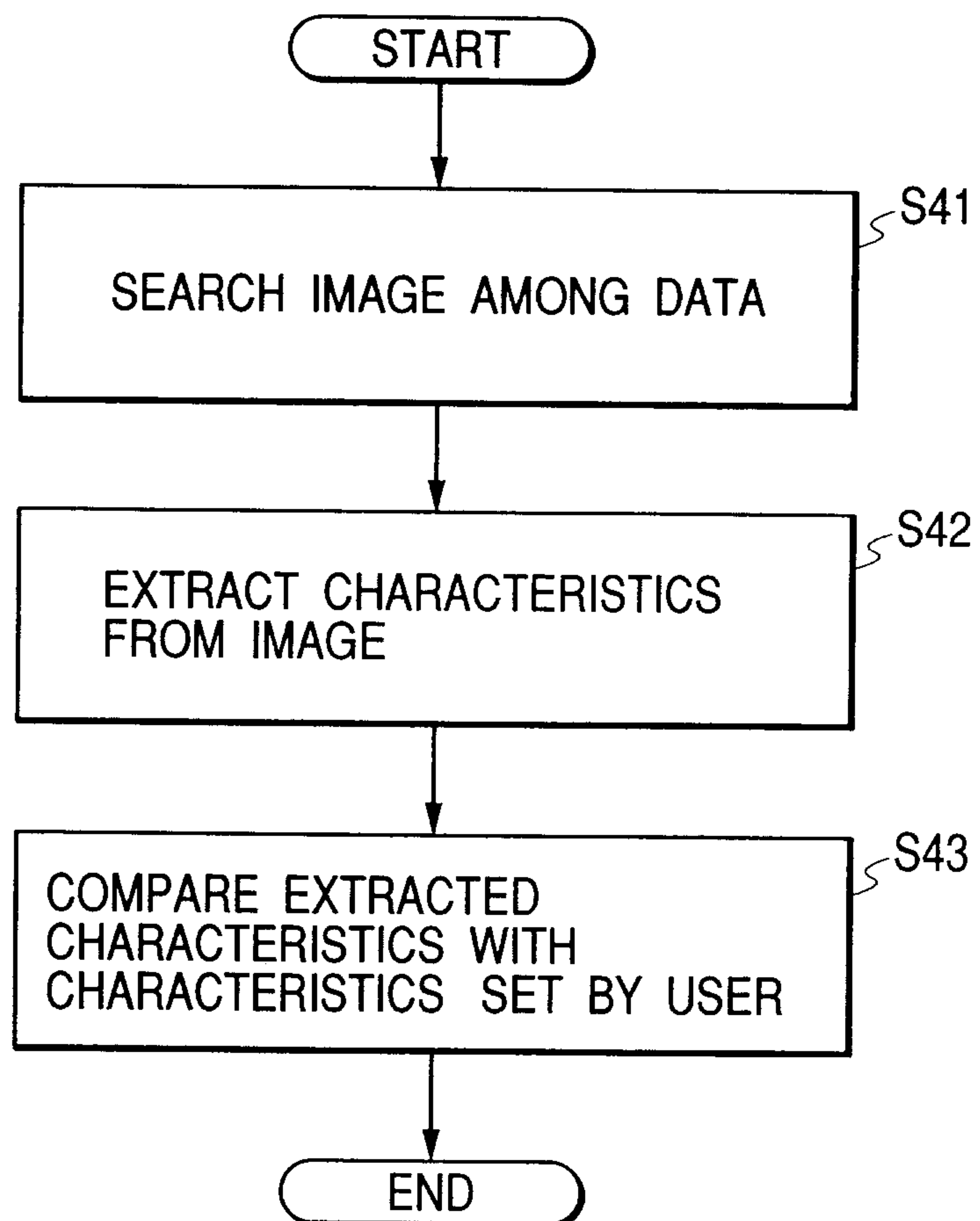
**FIG. 11**

11 / 13

**FIG. 12**



12 / 13

**FIG. 13**

**FIG. 14 PRIOR ART**