

SIMILARITY DETECTING APPARATUS  
AND DIRECTIONAL NEAREST NEIGHBOR DETECTING METHOD

ABSTRACT

5           In order to detect similar data from a great  
deal of data at high speed, a similarity detecting  
apparatus includes a random number generating unit 3  
which calculates random number data based on a  
parameter list, a table managing unit 5 which  
10 calculates a plurality of key calculation functions  
based on the random number data, calculates a  
plurality of tables based on input data, and  
calculates a candidate data list based on a query  
shown by a search condition, and a data processing  
15 unit 6 which calculates search result data from the  
candidacy data list to satisfy the condition shown a  
search condition. Each table is calculated for a  
value corresponding to a key to show a data list of  
data in which the data substituted into a key  
20 calculation function is equal to the key. A candidate  
data list contains a plurality of search data list and  
a search data list corresponding to the table shows  
the value corresponding to a query value obtained by  
substituting the query into the key calculation  
25 function.

Fig. 1