CLAIMS

[CLAIM 1] A personal safety verification system comprising:

a portal server that provides a database in which a plurality of sets, each including a feature value of a face of a missing person, personal information of a verification requesting person or the missing person, and a point of contact of the verification requesting person, are accumulated; and

a field server that acquires a copy of the database, extracts a feature value from a face image of a person captured by a camera, and searches for a similar feature value from the copy, wherein

the portal server accumulates and provides the feature value of the face in a state of being applied with template protection, and

when the field server succeeds in searching of the similar feature value, the field server presents the personal information corresponding to the similar feature value to the captured person and asks for verification, and when accepting a verification operation from the person, notifies the point of contact of the verification requesting person.

[CLAIM 2] The personal safety verification system according to claim 1, wherein

the personal information is a telephone number,

the point of contact of the verification requesting person is an email address or a telephone number, and

the feature value of the face is vector data of 100 dimensions or less or having a data size of 128 bytes or less, which is reduced in dimensions by main component analysis, independent component analysis, or linear discriminant analysis, and is applied with template protection by using at least one of random projection, a one-way function, and public key encryption.

[CLAIM 3] The personal safety verification system according to claim 1, wherein

the portal server is connected with Internet and receives the set including the feature value, to which the template protection is applied, from a terminal of the verification requesting person, and when the field server has an electronic certificate or is reliable, or only in a case of emergency, the portal server provides the database, and

a notification to the point of contact of the verification requesting person includes information indicating a place where the person who is verified is captured.

[CLAIM 4] A similarity search method performed on data encrypted for confidentiality, the method comprising:

a first step of encrypting a plurality of units of sampled data for confidentiality by

mapping in which a distance is preserved at least locally in a sampled data space;

a second step of clustering and recording a pair of the sampled data encrypted for confidentiality and arbitrary data, based on the sampled data encrypted for confidentiality itself;

a third step of encrypting query data for confidentiality by a method same as the method used for the sampled data;

a fourth step of identifying a cluster in which sampled data similar to the query data is recorded, based on the query data encrypted for confidentiality; and

a fifth step of identifying, from the identified cluster, a unit of sampled data similar to the query data, by calculating a distance between the sampled data encrypted for confidentiality and the query data encrypted for confidentiality, and accessing arbitrary data paired with the unit of the sampled data.

[CLAIM 5] The similarity search method performed on the data encrypted for confidentiality, according to claim 4, wherein

the encrypting the data for confidentiality in the first step and the third step includes:

a sub step of quantizing the data on a data space thereof;

a sub step of subtracting the quantized data from the data to obtain a residual; and

a sub step of inputting the quantized data to a cryptographic hash function to obtain a hash value,

wherein a set of the hash value and the residual is output as data encrypted for confidentiality, and

the clustering in the second step includes, based on part of the hash value of the sampled data encrypted for confidentiality, associating the hash value and a cluster one to one or many to one.

[CLAIM 6] The similarity search method performed on the data encrypted for confidentiality, according to claim 4, wherein

the encrypting the data for confidentiality in the first step and the third step includes:

a sub step of generating a random projection (RP) matrix having a number of columns corresponding to a number of elements of the data, based on a given specific number; and

a sub step of supposing the data to be a column vector and multiplying the RP matrix from a top to obtain data encrypted for confidentiality, and

the data is mapped to a space of the data encrypted for confidentiality in which a

distance is preserved substantially in an entire vector space.

[CLAIM 7] The similarity search method performed on the data encrypted for confidentiality, according to claim 6, wherein

the RP matrix is a unitary matrix.

[CLAIM 8] The similarity search method performed on the data encrypted for confidentiality, according to claim 4, wherein

the encrypting the data for confidentiality in the first step and the third step includes:

a first sub step of supposing the data to be a column vector, and encoding the data such that a hamming length becomes shorter as a distance is shorter locally in a vector space thereof;

a second sub step of, based on the encoded data, generating a random projection (RP) matrix randomized under a predetermined rule; and

a third sub step of multiplying the data supposed to be a column vector by the RP matrix from a top to obtain data encrypted for confidentiality, and

the clustering in the second step is performed such that units of sampled data encrypted for confidentiality, having a short distance with each other in a space of the sampled data encrypted for confidentiality, are in a same cluster.

[CLAIM 9] The similarity search method performed on the data encrypted for confidentiality, according to claim 8, wherein

the encoding in the first sub step uses one of locality sensitive hashing, vector quantization, error correction code, distance preserving run length limited code, permutation code, distance preserving map code, rank modulation code, and Gray code, or a combination thereof.

[CLAIM 10] The similarity search method performed on the data encrypted for confidentiality, according to claim 4, wherein

the encrypting the sampled data for confidentiality in the first step includes:

a first sub step of dividing the sampled data into two by a method of separating a high-order bit and a low-order bit in binary representation of each element;

a second sub step of Gray-coding data of the high-order bit obtained through the division into two for each element;

a third sub step of generating a random number;

a fourth sub step of altering a part of the Gray code under a predetermined rule based on the random number;

a fifth sub step of altering data of the low-order bit obtained through the division into two, based on the random number;

a sixth sub step of supposing the altered data of the low-order bit to be a column vector, and multiplying a random projection (RP) matrix, generated based on the altered Gray code, from a top, to obtain data of the low-order bit to which random projection is applied;

a seventh sub step of applying error correction coding to the altered Gray code to obtain a redundant symbol; and

an eighth sub step of linking the data of the low-order bit, to which the random projection is applied, and the redundant symbol to obtain data encrypted for confidentiality.

[CLAIM 11] The similarity search method performed on the data encrypted for confidentiality, according to claim 4, wherein

in the encrypting the query data for confidentiality in the third step includes:

a first sub step of dividing the query data into two by a method of separating a high-order bit and a low-order bit in binary representation of each element;

a second sub step of Gray-coding data of the high-order bit obtained through the division into two for each element;

a third sub step of determining, from among numbers of entire sets having a possibility to be generated as random numbers, one on which a trial is to be performed;

a fourth sub step of altering a part of the Gray code under a predetermined rule, based on the one number on which the trial is to be performed;

a fifth sub step of applying error correction coding to the altered Gray code to obtain a redundant symbol;

a sixth sub step of applying error correction decoding to the Gray code before alteration by using the redundant symbol;

a seventh sub step of altering data of the low-order bit obtained by the division into two, based on a position of a bit corrected by the error correction decoding;

an eighth sub step of supposing data of the altered low-order bit to be a column vector, multiplying a random projection (RP) matrix, generated based on the altered Gray code, from a top to obtain data of the lower-order bit to which random projection is applied, and

a ninth sub step of linking the data of the low-order bit, to which the random projection is applied, and the redundant symbol to obtain data encrypted for confidentiality.

[CLAIM 12] The similarity search method performed on the data encrypted for confidentiality, according to claim 11, wherein

the error correction coding in the fifth sub step has an error correction capability

of not less than a number obtained by adding 1 to a double number of the number of bits having a possibility that the Gray code is altered, and

the identifying the cluster in the fourth step includes narrowing down clusters having a possibility that sampled data similar to the query data is recorded, to a number less than a number of times of trials in the third sub step.

[CLAIM 13] The similarity search method performed on the data encrypted for confidentiality, according to claim 4, wherein

the units of sampled data are an image feature value vector in which dimensionality is reduced by main component analysis, independent component analysis, or linear discriminant analysis, or biometrics information.

[CLAIM 14] A computer-readable tangible medium storing a program capable of causing a computer to execute:

a first step of encrypting a plurality of units of sampled data for confidentiality by mapping in which a distance is preserved at least locally in a sampled data space;

a second step of clustering and recording a pair of the sampled data encrypted for confidentiality and arbitrary data, based on the sampled data encrypted for confidentiality itself;

a third step of encrypting query data for confidentiality by a method same as the method used for the sampled data;

a fourth step of identifying a cluster in which sampled data similar to the query data is recorded, based on the query data encrypted for confidentiality; and

a fifth step of identifying, from the identified cluster, a unit of sampled data similar to the query data, by calculating a distance between the sampled data encrypted for confidentiality and the query data encrypted for confidentiality, and accessing arbitrary data paired with the unit of the sampled data.

CLAIMS

[CLAIM 1] A similarity search method performed on data encrypted for confidentiality, the method comprising:

a first step of encrypting a plurality of units of sampled data for confidentiality by mapping in which a distance is preserved at least locally in a sampled data space;

a second step of clustering and recording a pair of the sampled data encrypted for confidentiality and arbitrary data, based on the sampled data encrypted for confidentiality itself;

a third step of encrypting query data for confidentiality by a method same as the method used for the sampled data;

a fourth step of identifying a cluster in which sampled data similar to the query data is recorded, based on the query data encrypted for confidentiality; and

a fifth step of identifying, from the identified cluster, a unit of sampled data similar to the query data, by calculating a distance between the sampled data encrypted for confidentiality and the query data encrypted for confidentiality, and accessing arbitrary data paired with the unit of the sampled data,

wherein

the encrypting the data for confidentiality in the first step and the third step includes:

a sub step of quantizing the data on a data space thereof;

a sub step of subtracting the quantized data from the data to obtain a residual; and

a sub step of inputting the quantized data to a cryptographic hash function to obtain a hash value,

wherein a set of the hash value and the residual is output as data encrypted for confidentiality, and

the clustering in the second step includes, based on part of the hash value of the sampled data encrypted for confidentiality, associating the hash value and a cluster one to one or many to one.

[CLAIM 2] The similarity search method performed on the data encrypted for confidentiality, according to claim 1, wherein

the encrypting the data for confidentiality in the first step and the third step includes:

a sub step of generating a random projection (RP) matrix having a number of columns corresponding to a number of elements of the data, based on a given specific number; and

a sub step of supposing the data to be a column vector and multiplying the RP matrix from a top to obtain data encrypted for confidentiality, and

the data is mapped to a space of the data encrypted for confidentiality in which a distance is preserved substantially in an entire vector space.

[CLAIM 3] The similarity search method performed on the data encrypted for confidentiality, according to claim 2, wherein

the RP matrix is a unitary matrix.

[CLAIM 4] The similarity search method performed on the data encrypted for confidentiality, according to claim 1, wherein

the encrypting the data for confidentiality in the first step and the third step includes:

a first sub step of supposing the data to be a column vector, and encoding the data such that a hamming length becomes shorter as a distance is shorter locally in a vector space thereof;

a second sub step of, based on the encoded data, generating a random projection (RP) matrix randomized under a predetermined rule; and

a third sub step of multiplying the data supposed to be a column vector by the RP matrix from a top to obtain data encrypted for confidentiality, and

the clustering in the second step is performed such that units of sampled data encrypted for confidentiality, having a short distance with each other in a space of the sampled data encrypted for confidentiality, are in a same cluster.

[CLAIM 5] The similarity search method performed on the data encrypted for confidentiality, according to claim 4, wherein

the encoding in the first sub step uses one of locality sensitive hashing, vector quantization, error correction code, distance preserving run length limited code, permutation code, distance preserving map code, rank modulation code, and Gray code, or a combination thereof.

[CLAIM 6] The similarity search method performed on the data encrypted for confidentiality, according to claim 1, wherein

the encrypting the sampled data for confidentiality in the first step includes:

a first sub step of dividing the sampled data into two by a method of separating a high-order bit and a low-order bit in binary representation of each element;

a second sub step of Gray-coding data of the high-order bit obtained through the division into two for each element;

a third sub step of generating a random number;

a fourth sub step of altering a part of the Gray code under a predetermined rule based on the random number;

a fifth sub step of altering data of the low-order bit obtained through the division into two, based on the random number;

a sixth sub step of supposing the altered data of the low-order bit to be a column vector, and multiplying a random projection (RP) matrix, generated based on the altered Gray code, from a top, to obtain data of the low-order bit to which random projection is applied;

a seventh sub step of applying error correction coding to the altered Gray code to obtain a redundant symbol; and

an eighth sub step of linking the data of the low-order bit, to which the random projection is applied, and the redundant symbol to obtain data encrypted for confidentiality.

[CLAIM 7] The similarity search method performed on the data encrypted for confidentiality, according to claim 1, wherein

in the encrypting the query data for confidentiality in the third step includes:

a first sub step of dividing the query data into two by a method of separating a high-order bit and a low-order bit in binary representation of each element;

a second sub step of Gray-coding data of the high-order bit obtained through the division into two for each element;

a third sub step of determining, from among numbers of entire sets having a possibility to be generated as random numbers, one on which a trial is to be performed;

a fourth sub step of altering a part of the Gray code under a predetermined rule, based on the one number on which the trial is to be performed;

a fifth sub step of applying error correction coding to the altered Gray code to obtain a redundant symbol;

a sixth sub step of applying error correction decoding to the Gray code before alteration by using the redundant symbol;

a seventh sub step of altering data of the low-order bit obtained by the division into two, based on a position of a bit corrected by the error correction decoding;

an eighth sub step of supposing data of the altered low-order bit to be a column vector, multiplying a random projection (RP) matrix, generated based on the altered Gray code, from a top to obtain data of the lower-order bit to which random projection is applied, and

a ninth sub step of linking the data of the low-order bit, to which the random projection is applied, and the redundant symbol to obtain data encrypted for confidentiality.

[CLAIM 8] The similarity search method performed on the data encrypted for confidentiality, according to claim 7, wherein

the error correction coding in the fifth sub step has an error correction capability of not less than a number obtained by adding 1 to a double number of the number of bits having a possibility that the Gray code is altered, and

the identifying the cluster in the fourth step includes narrowing down clusters having a possibility that sampled data similar to the query data is recorded, to a number less than a number of times of trials in the third sub step.

[CLAIM 9] The similarity search method performed on the data encrypted for confidentiality, according to claim 1, wherein

the units of sampled data are an image feature value vector in which dimensionality is reduced by main component analysis, independent component analysis, or linear discriminant analysis, or biometrics information.

[CLAIM 10] A personal safety verification system comprising:

a portal server that provides a database in which a plurality of sets, each including a feature value of a face of a missing person, personal information of a verification requesting person or the missing person, and a point of contact of the verification requesting person, are accumulated; and

a field server that acquires a copy of the database, extracts a feature value from a face image of a person captured by a camera, and searches for a similar feature value from the copy, wherein

the portal server accumulates and provides the feature value of the face in a state of being applied with template protection, and

when the field server succeeds in searching of the similar feature value, the field server presents the personal information corresponding to the similar feature value to the captured person and asks for verification, and when accepting a verification operation from the person, notifies the point of contact of the verification requesting person.

[CLAIM 11] The personal safety verification system according to claim 10, wherein

the personal information is a telephone number,

the point of contact of the verification requesting person is an email address or a telephone number, and

the feature value of the face is vector data of 100 dimensions or less or having a data size of 128 bytes or less, which is reduced in dimensions by main component analysis, independent component analysis, or linear discriminant analysis, and is applied with template

protection by using at least one of random projection, a one-way function, and public key encryption.

[CLAIM 12] The personal safety verification system according to claim 10, wherein

the portal server is connected with Internet and receives the set including the feature value, to which the template protection is applied, from a terminal of the verification requesting person, and when the field server has an electronic certificate or is reliable, or only in a case of emergency, the portal server provides the database, and

a notification to the point of contact of the verification requesting person includes information indicating a place where the person who is verified is captured.