Office de la propriété intellectuelle du Canada
Un organisme d'Industrie Canada
www.opic.gc.ca

Canadian Intellectual Property Office
An Agency of Industry Canada
www.cipo.gc.ca

June 19, 2013

WEININGER, ARTHUR M.
200 Rosedale Heights Dr.
THORNHILL Ontario
L4J 4W4

Application No.	:	**2,818,017**
Owner	:	WEININGER, ARTHUR M., WEININGER, SUSAN
Title	:	**N-DIMENSIONAL DATA PIPELINE PROCESSOR HAVING ADVANCED FEATURES**

Dear Madam / /Sir:

Your application for Patent has been allotted the above mentioned serial number and filing date of **June 11, 2013.**

When corresponding with the Office, please refer to the above referenced application number and the name of the applicant or inventor.

For General enquiries, please call our toll-free number at 1-866-997-1936. You can also consult the guide to patent at the following web address: http://www.cipo.ic.gc.ca/eic/site/cipointernet-internetopic.nsf/eng/wr00102.html or contact the undersigned.

Lyne Éthier
Patent Formalities and Assignments Analyst
819-953-7554

Office de la propriété intellectuelle du Canada
Un organisme d'Industrie Canada
www.opic.gc.ca

Canadian Intellectual Property Office
An Agency of Industry Canada
www.cipo.gc.ca

WEININGER, ARTHUR M.
200 Rosedale Heights Dr.
THORNHILL Ontario
L4J 4W4

Date : 2013/06/19

FILING CERTIFICATE

Application No. : **2,818,017**
Filing Date : 2013/06/11

Expected Open-to-Public Inspection Date : 2014/12/11
Your Reference :

Title of Invention : N-DIMENSIONAL DATA PIPELINE PROCESSOR HAVING ADVANCED FEATURES

Applicant(s) : WEININGER, ARTHUR M.; WEININGER, SUSAN

Inventor(s) : WEININGER, ARTHUR M.; WEININGER, SUSAN

Special Notice

You are reminded that annual fees to maintain your application are needed for each one-year period between the 2nd and 20th anniversaries and must be paid on or before each anniversary. Failure to pay within the prescribed time limit will lead to abandonment of your application.

Commissioner of Patents

Canada

OPIC CIPO

N-DIMENSIONAL DATA PIPELINE PROCESSOR HAVING ADVANCED FEATURES

ABSTRACT

An n-dimensional data pipeline processor and method are provided. This method is useful for the efficient characterization of spatial occupancy and identification of common spatial occupancy. The n-dimensional data pipeline processor data input preparing apparatus processes n-dimensional data by parsing, attributing, shifting, sizing, and truncating the data according to data entry modifier instructions. The n-dimensional data pipeline processor matrix output preparing apparatus makes matrix indices that represent characterized spatial occupancy of processed data and, using data entry modifier instructions, makes matrix indices that represent common spatial occupancy in the processed data. The n-dimensional data pipeline processor feature output preparing apparatus uses the matrix indices that represent common spatial occupancy to select data elements from the input data that contribute elements to the identified common spatial occupancy. Output matrix indices and selected data provide feedback to the n-dimensional data pipeline processor.

Inventors: Weininger; Arthur (Thornhill, Ontario) and Weininger; Susan (Thornhill, Ontario)

References Cited

U.S. PATENT DOCUMENTS

7,808,503	10/2010	Duluk Jr., J.F. et al.
6,804,683	10/2004	Matsuzaki, H. et al.

OTHER PUBLICATIONS

Blaney, J.M.; Jorgensen, E.C.; Connolly, M.L.; Ferrin, T.E.; Langridge, R.; Oatley, S.J.; Burridge, J.M.; Blake, C.C.; "Computer graphics in drug design: molecular modeling of thyroid hormone–prealbumin interactions." *J. Med. Chem.* (1982) 25(7):785-790

Protein Data Bank ID: 1AIY.
Chang, X.; Jorgensen, A.M.; Bardrum, P.; Led, J.J.; "Solution structures of the R6 human insulin hexamer," *Biochemistry* (1997) 36: 9409-9422

Protein Data Bank ID: 1MEY.
Kim, C.A. and Berg, J.M.; "A 2.2 A resolution crystal structure of a designed zinc finger protein bound to DNA." *Nat. Struct. Biol.* (1996) 3: 940-945

Protein Data Bank ID: 3D1M.
McLellan, J.S.; Zheng, X.; Hauk, G.; Ghirlando, R.; Beachy, P.A.; Leahy, D.J.; "The mode of Hedgehog binding to Ihog homologues is not conserved across different phyla." *Nature* (2008) 455: 979-983

Protein Data Bank ID: 3M7K.
Shen, B.W.; Heiter D.F.; Chan, S.H.; Wang, H.; Xu S.Y.; Morgan, R.D.; Wilson, G.G.; Stoddard, B.L.; "Unusual target site disruption by the rare-cutting HNH restriction endonuclease PacI.", *Structure* (2010) 18: 734-743

N-DIMENSIONAL DATA PIPELINE PROCESSOR HAVING ADVANCED FEATURES

TECHNICAL FIELD OF THE INVENTION

The present invention relates to computing systems generally, to the pipeline processing of n-dimensional data more particularly, and even more particularly to a method implementing enhanced data preparing, processing, and feedback on the processing. Furthermore, the present invention relates to methods for preparing, comparing, and visualizing data, and more particularly to a method and an apparatus for characterizing spatial occupancy of n-dimensional data elements, identifying data elements that have common spatial occupancy, and feeding this information back into the n-dimensional data pipeline processor of the present invention for use in further processing.

BACKGROUND OF THE INVENTION

The spatial correspondence of elements of n-dimensional data sets is important in the evaluation of n-dimensional data sets. For example, the correlation of the fit of two molecular surfaces with binding affinities between two molecules is described in: Blaney, J.M.; Jorgensen, E.C.; Connolly, M.L.; Ferrin, T.E.; Langridge, R.; Oatley, S.J.; Burridge, J.M.; Blake, C.C. “Computer graphics in drug design: molecular modeling of thyroid hormone – prealbumin interaction. *J. Med. Chem.* (1982) 25(7):785-790, hereinafter referred to as the Blaney et al. reference. Blaney et al. assessed surfaces to determine the fit of a portion of the surface of prealbumin, thought to be the binding pocket for thyroxine, with each of the surfaces of thyroxine and its analogues. Thyroxine analogues with high binding affinity for prealbumin were observed to fit better in the binding pocket than low-affinity thyroxine analogues. The Blaney et al. reference suggests that this method would be a good screen for predicting the activity of untested thyroid hormones and accounts for the structure-activity relationship. One of the limitations with the Blaney et al. approach is that it draws conclusions about structure activity from a small number of static structures without

any scientific control set. The central observation of the Blaney et al. reference can neither be confirmed or discarded because it was not determined whether, in the known crystallographic database, there were other thyroxine analogue-binding protein sites that were different in shape from those identified or sites that were similar in shape but had no binding affinity. Methods that allow the exhaustive comparison of the fit between small molecules and their protein targets is necessary to test whether particular spatial configurations of atoms in molecules are related to their binding. An efficient n-dimensional pipeline processor that identified the features of potential binding sites would allow a proper scientific control to be run in structural studies of the fit between two molecules, whether the fit is assessed between the atoms of the molecules or surfaces of the atoms of the molecules.

In addition, in surface comparisons such as in the Blaney et al. reference, the number of operations to compare the points in two surfaces is not optimized. The comparison of how well two surfaces fit has historically required the assessment of the distances between every point in one set (e.g., the points representing the surface of the first molecule) to every point in the second set (e.g., the points representing the surface of the second molecule). Distance in common parlance is a scalar measure of how close or far apart points are to each other. In Euclidean geometry, distance is independent of the frame of reference of the points, so the distance from point A to point B is the same as the distance from point B to point A. For a point (x_1,x_2, …, x_n) and a point (y_1, y_2, …, y_n), distance is given by:

$$d = \left(\sum_{i=1}^{n} |x_i - y_i|^2 \right)^{\frac{1}{2}}$$

Point to point comparisons between two sets scale with the square of the number of point pairs whose distance must be computed. The time and computational requirements of an n-squared process are more onerous the higher the dimensionality of data and the larger the data set. In addition to an undesirable computational scale for surface comparision, there is no efficient mechanism in the method in Blaney et al. for identifying features of the data, in this case binding site configurations, and determining whether these features do exist or do not exist elsewhere in a larger data set.

The use of computer graphics to render and display objects is described in: US PATENT 6,804,683, herein referred to as the '683 patent. The '683 patent claims “a similar

image retrieving apparatus comprising: a region of interest setting portion for setting a portion of a three-dimensional image of retrieved object including a landmark as a region of interest; a feature quantity calculation portion for calculating a feature quantity including positions of the landmark and of the region of interest set by the region of interest setting portion as parameters; a similarity calculation portion for calculating a similarity between the retrieved object and three-dimensional image data stored in the image database in advance by comparing the feature quantity of the region of interest of the retrieved object which quantity is calculated with the feature quantity calculation portion with the three-dimensional image data where a region of interest stored in the image database is set in advance and the feature quantity of this region of interest is calculated; and an image selection portion for selecting similar images in an order of decreasing similarity from the image database on the basis of the similarity calculated with the similarity calculation portion."

The limitation of the '683 patent is that the "landmarks" must be predefined and the "region of interest" stored in the database must be set in advance. The "similarity calculation" is specifically characterized by calculating the similarity of the image with respect to only the image with which a key word agrees out of the key words selected from a set of key words prepared in advance. Images to which key words are added in advance are stored in the image database limiting how similarity is assessed. In the '683 patent, a three-dimensional standard model of an object which forms an object of comparison is prepared in advance, and the coordinate system particular to the three-dimensional standard model is defined in advance. The '683 patent is limited in that it requires a three-dimensional standard model of an object which forms an object of comparison in advance.

The use of computer graphics to render and display objects is described in: US PATENT 7,808,503, herein referred to as the '503 patent. The '503 patent describes methods for rendering objects and displaying the objects on a computer graphics device. The '503 patent invention claims include a deferred graphics pipeline processor comprising: "a geometry apparatus configured to receive primitive data related to a vertex on a surface and output a data stream in response thereto; a mode extraction apparatus configured to receive data stream from the geometry apparatus and separate data stream into spatial data and non-spatial data; a sorting apparatus configured to receive the spatial data from the mode

extraction apparatus for storage; a multi-buffered polygon memory configured to receive the non-spatial data from the mode extraction apparatus for storage; and a mode injection apparatus configured to retrieve at least a portion of the non-spatial data from the polygon memory and output retrieved non-spatial data; wherein the mode injection apparatus is associated with at least one cache to determine whether the retrieved non-spatial data is cached, and the mode injection apparatus is operative to read the non-spatial data previously stored in a first frame simultaneously while the mode extraction apparatus is operative to store the non-spatial data in a second frame." A limitation of the '503 patent is that the input data and non-spatial data mapping are limited to to the vertices and polygons that define the spatial data structures. The '503 patent does not teach an efficient mechanism for identifying common spatial occupancy of n-dimensional data elements or disclose a means to figure out the inherent frame of common spatial reference between data sets.

BRIEF SUMMARY OF THE INVENTION

An object of the present invention is to provide an n-dimensional data pipeline processor that inputs one or more of any type of n-dimensional data set, renders information about characterized spatial occupancy and common spatial occupancy of n-dimensional data elements, and provides a mechanism for feedback of this information to additional cycles of data processing. Embodiments of the n-dimensional pipeline processor and method may include one or more of: data input preparing, matrix output preparing, and feature output preparing. Examples of n-dimensional databases include, but are not limited to: molecular composition and structure data, e.g., x-ray diffraction and NMR-derived structual data; financial data, image data, e.g., video, sound, or geophysical data; and symbolic data, e.g., text or numeric data. FIG. 1 is a block diagram explaining the structure of the n-dimensional data pipeline processor. NDDPP **(1)** accepts data of any dimensionality and efficiently evaluates input data for common spatial occupancy of n-dimensional data elements. NDDPP **(1)** comprises: the data input preparing apparatus, DIPA **(2)**; the matrix output preparing apparatus, MOPA **(3)**; and the feature output preparing apparatus, FOPA **(4)**.

As illustrated in FIG. 1, USER **(5)** initially outputs one data set, DS1 **(6)**, or more data

sets, DS2 **(7)**, and data entry modifier instructions, DEM **(8)**, into the NDDPP **(1)**. USER **(5)** may be any entity or more than one entity that can communicate information to and receive information from the NDDPP **(1)**, including devices, objects, and organisms. One skilled in the art will understand that USER **(5)** communication with NDDPP **(1)** can be accomplished by many and varied means, including, but not limited to: laser modulation to or from a receiver; communication with a data line, computer, or phone; direct sensory input to or from a biological organism or device. As illustrated in FIG. 1, USER **(5)** receives input data from NDDPP **(1)** in the form of M1 **(11)**, M2 **(12)**, and MC **(13)** from MOPA **(3)**, and DSC **(14)**, from FOPA **(4)**. USER **(5)** may, among other options, use, display, monitor, communicate, ignore, change, add to, delete, integrate, accept, and/or reject data coming from the NDDPP **(1)**.

As illustrated by FIG. 1, n-dimensional data, DS1 **(6)** and DS2 **(7)**, if any, is input to the DIPA **(2)** of NDDPP **(1)** from USER **(5)**. USER **(5)** also enters data entry modifier instructions, DEM **(8)**. If DSC **(14)** has been produced from FOPA **(4)**, DSC **(14)** would also be input to DIPA **(2)**. DIPA **(2)** parses and attributes the input data set or data sets. All parsed and attributed data is shifted, sized, and truncated by DIPA **(2)** to make parsed, attributed, shifted, sized, and truncated data PDS1 **(9)** and PDS2 **(10)**. PDS1 **(9)** represents the parsed, attributed, shifted, sized, and truncated data derived from DS1 **(6)**. PDS2 **(10)** represents the parsed, attributed, shifted, sized, and truncated data derived from data other than DS1 **(6)** such as DS2 **(7)**. PDS1 **(9)** and PDS2 **(10)** are output to MOPA **(3)**.

As illustrated by FIG. 1, PDS1 **(9)** and PDS2 **(10)** are input to MOPA **(3)** where matrix indices M1 **(11)** and M2 **(12)** are made that represent characterized spatial occupancy of n-dimensional data elements. M1 **(11)** represents the matrix indices made from PDS1 **(9)**. M2 **(12)** represents the matrix indices made from PDS2 **(10)**. Additional matrix indices MC **(13)** that represent common spatial occupancy of n-dimensional data elements are then made using M1 **(11)** and M2 **(12)**. M1 **(11)**, M2 **(12)** and MC **(13)** are output to USER **(5)**. MC **(13)** is also output to FOPA **(4)** and MOPA **(3)**.

As illustrated by FIG. 1, DS1 **(6),** DS2 **(7),** and MC **(13)** are input to FOPA **(4)** where matrix indices MC **(13)** are used to select DS1 **(6)** and DS2 **(7)** data that contribute elements to the identified common spatial occupancy of n-dimensional data elements. The selected

data, DSC **(14)**, is then output to USER **(5)** and DIPA **(2)**.

DIPA **(2)** inputs n-dimensional data and, as instructed, parses, attributes, shifts, sizes, and truncates the data. DIPA **(2)** is comprised of processors DIPA-PA **(17)** and DIPA-SST **(20)**. As illustrated in FIG. 2, DIPA **(2)** inputs data sets DS1 **(6)**, DS2 **(7)**, and DSC **(14)**, if any, and DEM **(8)** instruction sets DEMI **(15)** and DEMS **(16)**. DIPA-PA **(17)** parses and attributes input data according to DEMI **(15)** and outputs parsed and attributed data PADS1 **(18)** and PADS2 **(19)** to DIPA-SST **(20)**. DIPA-SST **(20)** shifts, sizes, and truncates PADS1 **(18)** and PADS2 **(19)** data according to DEMS **(16)** and outputs PDS1 **(9)** and PDS2 **(10)** to MOPA **(3)**.

MOPA **(3)** builds matrix indices that represent characterized spatial occupancy and common spatial occupancy of n-dimensional data elements. MOPA **(3)** makes matrix indices that represent characterized spatial occupancy of n-dimensional data elements by using parsed, attributed, shifted, and truncated data input from DIPA **(2)** as matrix index values. MOPA **(3)** makes matrix indices that represent common spatial occupancy of n-dimensional data elements using DEM **(8)** instruction set DEMM **(21)** to direct the production of matrix indices that have shared matrix index values. MOPA **(3)** is comprised of processors MOPA-MM **(22)** and MOPA-CM **(23)**. As illustrated in FIG. 3, MOPA-MM **(22)** inputs PDS1 **(9)** and PDS2 **(10)** and outputs matrix indices M1 **(11)** and M2 **(12)**. Using DEM **(8)** instruction set DEMM **(21)**, M1 **(11)**, M2 **(12)**, and any previous MC **(13)**, are further processed by MOPA-CM **(23)** to make the current MC **(13)**. M1 **(11)**, M2 **(12)**, and MC **(13)** are output to USER **(5)**. MC **(13)** is also output to MOPA-CM **(23)** closing a loop within MOPA **(3)**. MC **(13)**, fed back to subsequent rounds of MOPA-CM **(23)** processing, is subject to further modification, redaction, or disregard by MOPA-CM **(23)** using DEMM **(21)**. MC **(13)** is also output to FOPA **(4)** as illustrated in FIG. 1 and FIG. 3.

FOPA **(4)** uses matrix indices that represent common spatial occupancy of n-dimensional data elements to generate selection criteria. This selection criteria is used to select n-dimensional data from the input data. The selected n-dimensional data are data elements that have been identified that have features with common spatial occupancy. FOPA **(4)** is comprised of processors FOPA-CC **(24)** and FOPA-AC **(26)**. As illustrated in FIG. 4, FOPA **(4)** inputs DS1 **(6)**, DS2 **(7)**, and MC **(13)**. MC **(13)** is processed by FOPA-CC **(24)** to

make selection criteria SCF **(25)**. FOPA-CC **(24)** outputs SCF **(25)** to FOPA-AC **(26)**. FOPA-AC **(26)** inputs and selects DS1 **(6)** and DS2 **(7)** data as determined by SCF **(25)** and outputs DSC **(14)**. DSC **(14)** is output to DIPA-PA **(17)**, closing a loop within the NDDPP **(1)**. DSC **(14)** is also output to USER **(5)**. USER **(5)** can be a closed loop or open loop depending on whether DEM **(8)** or the USER **(5)** generates additional DEM **(8)**.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated in and constitute a part of the specification, illustrate embodiments of the invention, and together with the general description given above and the detailed description of the preferred embodiments given below, serve to explain the principles of the invention.

FIG. 1 is a block diagram for explaining the n-dimensional data pipeline processor, NDDPP **(1)**.

FIG. 2 is a block diagram for explaining the data input preparing apparatus, DIPA **(2)**.

FIG. 3 is a block diagram for explaining the matrix output preparing apparatus, MOPA **(3)**.

FIG. 4 is a block diagram for explaining the feature output preparing apparatus, FOPA **(4)**.

FIG. 5 is a flowchart showing each step of the processing procedure of the data input preparing apparatus, DIPA **(2)**.

FIG. 6 is a flowchart showing each step of the processing procedure of the matrix output preparing apparatus, MOPA **(3)**.

FIG. 7 is a flowchart showing each step of the processing procedure of the feature output preparing apparatus, FOPA **(4)**.

FIG. 8 is a view showing a visualization of FOPA **(4)** output of an example of DSC **(14)**: zinc, from DS1-M, and carbon and nitrogen atoms, from DS2b-M, that share a common spatial occupancy of a volume.

FIG. 9 is a view showing a visualization of FOPA **(4)** output of an example of DSC **(14)**: zinc, three CE1 carbons, and three NE2 nitrogen atoms from 1AIY.PDB that share a common spatial occupancy of a volume.

FIG. 10 is a view showing a visualization of FOPA **(4)** output of an example of DSC **(14)**: three CE1 carbon atoms and three NE2 nitrogen atoms from 1AIY.PDB that share a common spatial occupancy of a volume.

FIG. 11 is a view showing a visualization of FOPA **(4)** output of an example of DSC **(14)**: a zinc and three histidine residues from 1AIY.PDB which contribute atoms to a common spatial occupancy of a volume.

FIG. 12 is a view showing a visualization of FOPA **(4)** output of an example of DSC **(14)**: residues from six chains from 1AIY.PDB in the insulin hexamer that contribute atoms to a common spatial occupancy of three phenol molecules.

FIG. 13 is a view showing a visualization of FOPA **(4)** output of an example of DSC **(14)**: four CE1 carbon atoms and four NE2 nitrogen atoms from 3D1M.PDB that share a common spatial occupancy of a volume with a zinc atom; and distances between two of the four NE2 nitrogen atoms to the zinc atom.

FIG. 14 is a view showing a visualization of FOPA **(4)** output of an example of DSC **(14)**: two SG sulfur atoms from cysteine residues and two NE2 atoms from histidine residues in a tetrahedral configuration from 1MEY.PDB that share a common spatial occupancy of a volume with a zinc atom.

FIG. 15 is a view showing a visualization of FOPA **(4)** output of an example of DSC **(14)**: two sets of four SG sulfur atoms from cysteine residues in tetrahedral configurations from 3M7K.PDB where each set of four sulfur atoms share a common spatial occupancy of a volume with a zinc atom.

FIG. 16 is a view showing a visualization of MOPA **(3)** output of an example of M1 **(11)** and M2 **(12)**: the characterized spatial occupancy of words with similar meaning.

FIG. 17 is a view showing a visualization of MOPA **(3)** output of an example of MC **(13):** the common spatial occupancy of words with similar meaning.

FIG. 18 is a view showing a visualization of DIPA **(2)** output of an example of PDS1 **(9)** and PDS2 **(10)**: characterized spatial occupancy in processed image sectors.

FIG. 19 is a view showing a visualization of MOPA **(3)** output of an example of MC **(13)**: matrix indices representing common spatial occupancy of data features in input image sectors.

DETAILED DESCRIPTION OF THE INVENTION

The invention will now be described in detail by way of illustrations, examples, and pseudocode. The pseudocode used is an example of computer instructions using universal computer language conventions. While the pseudocode employed in this description has been invented solely for the purposes of this description, it will be easily understandable by any computer programmer skilled in the art. It will be readily apparent to one skilled in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims. We first provide a top-level system architectural description.

N-dimensional data sets are often formatted as n-tuples of data and, for simple sets, are often listed in a column or vector format. Components of the vector are often listed in columns. N-dimensional data may also be in an unformatted stream. The dimension of a vector space is the number of vectors in any basis for the space, i.e., the number of coordinates necessary to specify any vector. This notion of dimension in a vector space, or the cardinality of a basis, is often referred to as the Hamel dimension or algebraic dimension to distinguish it from other notions of dimension. The Hausdorff dimension is defined for all metric spaces and, unlike the Hamel dimension, can also attain non-integer real values. There are more definitions of fractal dimensions that are used for highly irregular sets and have non-integer positive real values. Some complex networks are characterized by fractal dimensions. The concept of dimension can be generalized to include networks embedded in space. The fractal dimensions may be used to characterize their spatial constraints.

While the NDDPP **(1)** according to the present invention will be explained using examples of simple n-dimensional data as input, one skilled in the art will appreciate that the present invention may be applied to more complex dimensionality. Processing of n-dimensional data by the n-dimensional data pipeline processor generates: matrix indices that represent characterized spatial occupancy of input n-dimensional data elements, matrix indices that represent common spatial occupancy of input n-dimensional data elements, and selected n-dimensional data selected using selection criteria generated from matrix indices

representing common spatial occupancy of input n-dimensional data elements.

First Embodiment

In order to demonstrate the operation of the n-dimensional data pipeline processor of the present invention in the first embodiment, 12-dimensional data sets specifying atoms, attributes of atoms, and their atomic coordinates will be used as the first example of n-dimensional data input to the NDDPP **(1)**. The first example of data input to the NDDPP **(1)** is excerpted from the Protein Data Bank structure file, 1AIY, herein referred to as 1AIY.PDB. 1AIY.PDB contains 10 NMR structures for the insulin hexamer and its associated small molecules and atoms. A companion publication to the 1AIY.PDB structure is Chang, X., Jorgensen, A.M., Bardum, P. and Led, J.J. Biochemistry (1997) 36: 9409-9422. Data from the first structure of the published 1AIY.PDB have been excerpted, reformatted, and listed at the end of this specification in Tables 1, 2, and 3. Tables 1, 2, and 3 are prior art and are given for informational purposes only, to provide input data for the exemplification of the first embodiment. Tables 1, 2, and 3 consist of structural information listings of the human insulin A and B chains and an associated zinc ("ZN") atom, which have been stripped of associated hydrogen atoms and water molecules. Table 1 contains a ZN atom, is given as an example of DS1 **(6)**, and is hereinafter referred to as DS1-M1. Table 2, containing insulin chain A, is given as an example of DS2 **(7)**, and is hereinafter referred to as DS2a-M1. Table 3, containing insulin chain B, is given as an example of more than one DS2 **(7)**, and is hereinafter referred to as DS2b-M1.

FIG. 5 is a flowchart showing each step of the processing procedure of DIPA **(2)** and outlines how the example data is processed through DIPA **(2)** of the NDDPP **(1)** as illustrated in FIG. 1 and FIG. 2. The first step in processing, as described in FIG. 5 Step **(27)** is to input data sets, DS1 **(6)** and DS2 **(7),** to DIPA-PA **(17)**. For the purposes of a detailed discussion of the processing of data by the NDDPP **(1)** of the present invention, the input data sets, DS1-M1, DS2a-M1, and DS2b-M1, as listed in Tables 1, 2, and 3, are input to DIPA **(2)**. Also input to DIPA **(2)** are the DEM **(8)** instruction sets DEMI **(15)** and DEMS **(16)**. Since DSC **(14)** is output data of NDDPP **(1)** processing, no DSC **(14)** is input on a first processing cycle of data by NDDPP **(1)**.

As illustrated in FIG. 2, DEMI **(15)** is input to DIPA-PA **(17)** and provides instructions for the parsing and attribution of the input data sets. For the purposes of this example, DEMI-M1 will be used as an example of the DEMI **(15)**. DEMI **(15)** includes how data input to DIPA-PA **(17)** should be parsed, attributed, or ignored. Throughout the specification, data in columns are referred to by a "C" followed by the column number, e.g., column 7 is referred to as "C7." DEMI-M1 example instructions contain the following information for formatting DS1-M1, DS2a-M1, and DS2b-M1:

```
DEMI-M1:

FOR ALL DS1-M AND DS2-M ENTRIES:

DATA PARSING:
  +C1 +C2 +C3 +C4 +C5 +C6 +C7 +C8 +C9
  -C10 -C11 -C12
DATA SEPARATION:
  TABS AND LINE RETURNS
  C1 THROUGH C12:
  1 RECORD NAME (“ATOM” or “HETATM”)
  2 ATOM SERIAL NUMBER IN THE DATA
  3 ATOM NAME
  4 RESIDUE NAME
  5 CHAIN IDENTIFIER
  6 RESIDUE SEQUENCE NUMBER
  7 ORTHOGONAL COORDINATES - atoms's x position (angstroms)
  8 ORTHOGONAL COORDINATES - atoms's y position (angstroms)
  9 ORTHOGONAL COORDINATES - atoms's x position (angstroms)
  10 OCCUPANCY
  11 TEMPERATURE FACTOR
  12 ELEMENT SYMBOL

FOR C7, C8, AND C9, CALCULATE C7LOW, C8LOW, AND C9LOW.
```

In this example, DIPA-PA **(17)** first parses each line of DS1-M1, DS2a-M1 and DS2b-M1 as instructed by DEMI-M1 as described in FIG. 5 Step **(27)**. The first line in DS1-M1 is:

```
HETATM 4729  ZN  ZN • 1   0.004  -0.434  -5.873 1.00 0.00 ZN
```

Note that, for clarity, the character '•' has been substituted for the blank space in C5 specifying the chain identifier in the original 1AIY.PDB.

Using DEMI-M1, the first line in DS1-M1 can be interpreted as describing a zinc atom, or ZN, with an atom serial number of 4729 in the first residue in the " • " chain (i.e., with a space defining the chain identifier) with an x, y, z position of its atom center as 0.004, -0.434, and -5.873 in cartesian coordinates measured in angstroms with an occupancy of 1.00 and a temperature factor of 0.000 whose element description is that of zinc, or ZN. DEMI-M1 instructs DIPA-PA **(17)** to ignore C10, C11, and C12 of the input data and, as an example, processes the first processed entry in DS1-M1 as:

```
HETATM 4729  ZN  ZN • 1   0.004  -0.434  -5.873
```

Similarly, the entries in DS2a-M1 and DS2b-M1 are processed in the DIPA-PA **(17)**. Specifically, and in illustration, the first line in DS2a-M1 is:

```
ATOM      1   N GLY A 1  18.988  -5.893  -1.830 1.00 0.00  N
```

Using DEMI-M1, the first line in DS2a-M1, as an example, can be interpreted as describing a nitrogen atom, or N, with an atom serial number of 1 in the first residue in the A chain which is a glycine, or GLY, with an x, y, z position of its atom center as 18.988, -5.893,-1.830 in cartesian coordinates measured in angstroms with an occupancy of 1.00 and a temperature factor of 0.000 whose element description is that of nitrogen, or N. For this example, DEMI-M1 instructs DIPA-PA **(17)** to ignore C10, C11, and C12. The first entry in DS2a-M1 is processed as:

```
ATOM      1   N GLY A 1  18.988  -5.893  -1.830
```

Attributes can either be added to data entries, substituted for data entries, or eliminated as data entries as instructed by DEMI-M1. In this example, the example DEMI-M1 instructs the DIPA-PA **(17)** to substitute the mapped values into C1, C3, C4, and C5:

```
DEMI-M1 (continued)

DATA ATTRIBUTES:
  C1 TO C1, C3 TO C3, C4 to C4, C5 TO C5
  IF C1 ENTRY EQUAL “ATOM”:    PUT  “1” IN C1
  IF C1 ENTRY EQUAL “HETATM”:  PUT  “2” IN C1
  IF C3 ENTRY EQUALS “C”:      PUT  “1” IN C3
  IF C3 ENTRY EQUALS “CA”:     PUT  “2” IN C3
  IF C3 ENTRY EQUALS “CB”:     PUT  “3” IN C3
  IF C3 ENTRY EQUALS “CD”:     PUT  “4” IN C3
  IF C3 ENTRY EQUALS “CD1”:    PUT  “5” IN C3
  IF C3 ENTRY EQUALS “CD2”:    PUT  “6” IN C3
  IF C3 ENTRY EQUALS “CE”:     PUT  “7” IN C3
  IF C3 ENTRY EQUALS “CE1”:    PUT  “8” IN C3
  IF C3 ENTRY EQUALS “CE2”:    PUT  “9” IN C3
  IF C3 ENTRY EQUALS “CG”:     PUT “10” IN C3
  IF C3 ENTRY EQUALS “CG1”:    PUT “11” IN C3
  IF C3 ENTRY EQUALS “CG2”:    PUT “12” IN C3
  IF C3 ENTRY EQUALS “CZ”:     PUT “13” IN C3
  IF C3 ENTRY EQUALS “N”:      PUT “14” IN C3
  IF C3 ENTRY EQUALS “ND1”:    PUT “15” IN C3
  IF C3 ENTRY EQUALS “ND2”:    PUT “16” IN C3
  IF C3 ENTRY EQUALS “NE”:     PUT “17” IN C3
  IF C3 ENTRY EQUALS “NE2”:    PUT “18” IN C3
  IF C3 ENTRY EQUALS “NH1”:    PUT “19” IN C3
  IF C3 ENTRY EQUALS “NH2”:    PUT “20” IN C3
  IF C3 ENTRY EQUALS “NZ”:     PUT “21” IN C3
  IF C3 ENTRY EQUALS “O”:      PUT “22” IN C3
```

```
IF C3 ENTRY EQUALS “OD1”:    PUT “23” IN C3
IF C3 ENTRY EQUALS “OE1”:    PUT “24” IN C3
IF C3 ENTRY EQUALS “OE2”:    PUT “25” IN C3
IF C3 ENTRY EQUALS “OG”:     PUT “26” IN C3
IF C3 ENTRY EQUALS “OG1”:    PUT “27” IN C3
IF C3 ENTRY EQUALS “OH”:     PUT “28” IN C3
IF C3 ENTRY EQUALS “OXT”:    PUT “29” IN C3
IF C3 ENTRY EQUALS “SG”:     PUT “30” IN C3
IF C3 ENTRY EQUALS “ZN”:     PUT “31” IN C3
IF C3 ENTRY EQUALS “C1”:     PUT “32” IN C3
IF C3 ENTRY EQUALS “C2”:     PUT “33” IN C3
IF C3 ENTRY EQUALS “C3”:     PUT “34” IN C3
IF C3 ENTRY EQUALS “C4”:     PUT “35” IN C3
IF C3 ENTRY EQUALS “C5”:     PUT “36” IN C3
IF C3 ENTRY EQUALS “C6”:     PUT “37” IN C3
IF C3 ENTRY EQUALS “O1”:     PUT “38” IN C3
IF C4 ENTRY EQUALS “ALA”:    PUT  “1” IN C4
IF C4 ENTRY EQUALS “ARG”:    PUT  “2” IN C4
IF C4 ENTRY EQUALS “ASN”:    PUT  “3” IN C4
IF C4 ENTRY EQUALS “ASP”:    PUT  “4” IN C4
IF C4 ENTRY EQUALS “CYS”:    PUT  “5” IN C4
IF C4 ENTRY EQUALS “GLN”:    PUT  “6” IN C4
IF C4 ENTRY EQUALS “GLU”:    PUT  “7” IN C4
IF C4 ENTRY EQUALS “GLY”:    PUT  “8” IN C4
IF C4 ENTRY EQUALS “HIS”:    PUT  “9” IN C4
IF C4 ENTRY EQUALS “ILE”:    PUT “10” IN C4
IF C4 ENTRY EQUALS “LEU”:    PUT “11” IN C4
IF C4 ENTRY EQUALS “LYS”:    PUT “12” IN C4
IF C4 ENTRY EQUALS “MET”:    PUT “13” IN C4
IF C4 ENTRY EQUALS “PHE”:    PUT “14” IN C4
IF C4 ENTRY EQUALS “PRO”:    PUT “15” IN C4
IF C4 ENTRY EQUALS “SER”:    PUT “16” IN C4
IF C4 ENTRY EQUALS “THR”:    PUT “17” IN C4
IF C4 ENTRY EQUALS “TRP”:    PUT “18” IN C4
IF C4 ENTRY EQUALS “TYR”:    PUT “19” IN C4
```

```
IF C4 ENTRY EQUALS “VAL”:    PUT “20” IN C4
IF C4 ENTRY EQUALS “ZN”:     PUT “21” IN C4
IF C4 ENTRY EQUALS “IPH”:    PUT “22” IN C4
IF C5 ENTRY EQUALS “A”:      PUT  “1” IN C5
IF C5 ENTRY EQUALS “B”:      PUT  “2” IN C5
IF C5 ENTRY EQUALS “•”:      PUT  “3” IN C5
IF C5 ENTRY EQUALS “C”:      PUT  “4” IN C5
IF C5 ENTRY EQUALS “D”:      PUT  “5” IN C5
IF C5 ENTRY EQUALS “E”:      PUT  “6” IN C5
IF C5 ENTRY EQUALS “F”:      PUT  “7” IN C5
IF C5 ENTRY EQUALS “G”:      PUT  “8” IN C5
IF C5 ENTRY EQUALS “H”:      PUT  “9” IN C5
IF C5 ENTRY EQUALS “I”:      PUT “10” IN C5
IF C5 ENTRY EQUALS “J”:      PUT “11” IN C5
IF C5 ENTRY EQUALS “K”:      PUT “12” IN C5
IF C5 ENTRY EQUALS “L”:      PUT “13” IN C5
```

DEMI **(15)** can also simply generate a set of attributes or assignments in order of their unique occurrence. For example, DEMI **(15)** could instruct DIPA-PA **(17)** to assign a number to every unique entry in the input data sets for any column of data. The same number of unique numbers would be assigned in each of C2 through C5 in an alternative example of this first embodiment: C2 would have two unique entries, C3 would have thirty-eight unique entries, C4 would have twenty-two unique entries, and C5 would have three unique entries. The number assignments in this alternative example would be listed in order of appearance in the data. Assignment by order of appearance for unique entries provides flexibility in that new numbers can be automatically added as new types of entries are processed without requiring a predetermination of the entry type. Direct assignment of attributes or unique numbers, as in the current example, provides the opportunity to group entries through choice of the assignment of number. Either method, applied to this example, would produce a unique value to be assigned for the data entry in the column.

As illustrated in FIG. 2, after processing by the DIPA-PA **(17)** as described in Step **(28)** of FIG. 5, PADS1 **(18)** and PADS2 **(19)** are made. For the purposes of this example,

PADS1-M will be an example of PADS1 **(18)** and PADS2a-M1 and PADS2b-M1 will be an example of PADS2 **(19)**. PADS1 **(18)** and PADS2 **(19)** are made after DIPA-PA **(17)** executes the example DEMI-M1. The PADS1-M1 first line, derived from DS1-M1 first line, the PADS2a-M1 first line, derived from DS2a-M1 first line, and the PADS2b-M first line, derived from DS2b-M1 first line, are respectively listed as follows:

```
2    4729  31  21  3  1   0.004  -0.434  -5.873
1       1  14   8  1  1  18.988  -5.893  -1.830
1     314  14  14  2  1   0.503   -8.36   6.002
```

In this example of the first embodiment, the parsed and attributed information, PADS1-M1, PADS2a-M1, and PADS2b-M1 are input to DIPA-SST **(20)** as illustrated in FIG. 2 and processed further according to instructions from DEMS **(16)** described in FIG. 5 Step **(29)**. In this example, the DIPA-SST **(20)** shifts, sizes, and truncates selected columns of PADS1-M1, PADS2a-M1, and PADS2b-M1 as instructed by the following example of a DEMS **(16)**, hereinafter referred to as DEMS-M1:

```
DEMS-M1:

FOR ALL PADS1-M1 AND PADS2-M1 DATA ENTRIES:

C1 TO C1, C2 TO C2, C3 TO C3, C4 TO C4, C5 TO C5, C6 TO C6

MULTFACTOR = 0.2

SHIFT COMMANDS:
  IF THE SMALLEST VALUE OF C7(“C7LOW”)IS NEGATIVE, THEN
    SUBTRACT THE VALUE OF C7LOW FROM EVERY C7 ENTRY AND
    PUT THE RESULT IN C7
  ELSE
    PUT C7 IN C7
  IF THE SMALLEST VALUE OF C8(“C8LOW”)IS NEGATIVE, THEN
```

```
      SUBTRACT THE VALUE OF C8LOW FROM EVERY C8 ENTRY AND
      PUT THE RESULT IN C8
    ELSE
      PUT C8 IN C8
    IF THE SMALLEST VALUE OF C9("C9LOW")IS NEGATIVE, THEN
      SUBTRACT THE VALUE OF C9LOW FROM EVERY C9 ENTRY AND
      PUT THE RESULT IN C9.
    ELSE
      PUT C9 IN C9

SIZE COMMANDS:
  MULTIPLY THE CURRENT VALUE OF C7 BY MULTFACTOR; AND
    PUT THE RESULT IN C7.
  MULTIPLY THE CURRENT VALUE OF C8 BY MULTFACTOR; AND
    PUT THE RESULT IN C8.
  MULTIPLY THE CURRENT VALUE OF C9 BY MULTFACTOR; AND
    PUT THE RESULT IN C9.

TRUNCATE COMMANDS:
  TRUNCATE THE CURRENT VALUE OF C7 AFTER THE DECIMAL PLACE;
    AND PUT THE INTEGER RESULT IN C7.
  TRUNCATE THE CURRENT VALUE OF C8 AFTER THE DECIMAL PLACE;
    AND PUT THE INTEGER RESULT IN C8.
  TRUNCATE THE CURRENT VALUE OF C9 AFTER THE DECIMAL PLACE;
    AND PUT THE INTEGER RESULT IN C9.
```

In this example of the first embodiment, an inspection of Tables 1, 2 and 3 show that the lowest and highest values for x, y, and z for the combined datasets DS1-M1, DS2a-M1, and DS2b-M1 are:

```
C7LOW =  -1.836
C8LOW = -18.712
C9LOW = -27.529
```

Therefore, in this example, as instructed by DEMS-M1, each line entry in PADS1-M1, PADS2a-M1 and PADS2b-M1 is processed:

For each number in C7: as instructed specifically by the SHIFT COMMAND, 1.836 is added to each number; as instructed specifically by the MULTIPLY COMMAND, the shifted result is multiplied by a multiplication factor, MULTFACTOR, of 0.2; as instructed specifically by the TRUNCATE COMMAND, the shifted and sized number is truncated after the decimal point; and the resultant integer value replaces the initial C7 number in C7.

For each number in C8: as instructed specifically by the SHIFT COMMAND, 18.712 is added to each number; as instructed specifically by the MULTIPLY COMMAND, the shifted result is multiplied by a multiplication factor, MULTFACTOR, of 0.2; as instructed specifically by the TRUNCATE COMMAND, the shifted and sized number is truncated after the decimal point; and the resultant integer value replaces the initial C8 number in C8.

For each number in C9: as instructed specifically by the SHIFT COMMAND, 27.529 is added to each number; as instructed specifically by the MULTIPLY COMMAND, the shifted result is multiplied by a multiplication factor, MULTFACTOR, of 0.2; as instructed specifically by the TRUNCATE COMMAND, the shifted and sized number is truncated after the decimal point; and the resultant integer value replaces the initial C9 number in C9.

To summarize DIPA **(2)** processing of DS1-M1, DS2a-M1, and DS2b-M1, the processing results of the first lines of each data set are reproduced for each step in FIG. 5.

After FIG. 5 Step **(27)**:

```
HETATM 4729  ZN  ZN • 1   0.004  -0.434  -5.873 1.00 0.00 ZN
ATOM      1   N GLY A 1  18.988  -5.893  -1.830 1.00 0.00  N
ATOM    314   N PHE B 1   0.503  -8.361   6.002 1.00 0.00  N
```

After FIG. 5 Step **(28)** PARSE:

```
HETATM 4729  ZN  ZN • 1   0.004  -0.434  -5.873
ATOM      1   N GLY A 1  18.988  -5.893  -1.830
ATOM    314   N PHE B 1   0.503  -8.361   6.002
```

After FIG. 5 Step **(28)** ATTRIBUTE:

```
2   4729  31  21  3  1   0.004  -0.434  -5.873
1      1  14   8  1  1  18.988  -5.893  -1.830
1    314  14  14  2  1   0.503  -8.361   6.002
```

After FIG. 5 Step **(29)** SHIFT:

```
2   4729  31  21  3  1   1.840  18.278  21.656
1      1  14   8  1  1  20.824  12.819  25.699
1    314  14  14  2  1   2.339  10.351  33.531
```

After FIG. 5 Step **(29)** MULTIPLY:

```
2   4729  31  21  3  1   0.368   3.656   4.331
1      1  14   8  1  1   4.165   2.564   5.140
1    314  14  14  2  1   0.468   2.070   6.706
```

After FIG. 5 Step **(29)** TRUNCATE and output of Step **(30)**:

```
2   4729  31  21  3  1      0      3      4
1      1  14   8  1  1      4      2      5
1    314  14  14  2  1      0      2      6
```

As described in FIG. 5 Step **(30)**, processed DS1-M1, DS2a-M1, and DS2b-M1 are output by DIPA **(2)** to MOPA **(3)** as PDS1 **(9)** and PDS2 **(10)**. In this example, PDS1 **(9)** derived from DS1-M1 will be referred to as PDS1-M1 and PDS2 **(10)** derived from DS2a-M1 and DS2b-M1 will be referred to as PDS2a-M1 and PDS2b-M1, respectively.

One skilled in the art will recognize that it is desirable, but not necessary, that SHIFT instructions from DEMS **(16)** produce positive column values in PDS1 **(9)** and PDS2 **(10)**. Negative values in PDS1 **(9)** and PDS2 **(10)** must be interpreted to be consistent with the spatial representation of positive values. One skilled in the art will recognize that DEMS **(16)** may recursively and/or incrementally adjust individual values given in a SHIFT instruction

and that a value of zero, or NO SHIFT, can be applied in any SHIFT instruction. One skilled in the art will also recognize that any number can be used as MULTFACTOR in DEMS **(16)**, including 1, and that different MULTFACTOR values may be applied to different data elements. For example, one column could be multiplied by 0.21, another by 0.30, and still another by 1.05. One skilled in the art will recognize that DEMS **(16)** may recursively and/or incrementally adjust individual values given in a MULTIPLY instruction.

FIG. 6 is a flowchart showing each step of the processing procedure of MOPA **(3)** and outlines how the example data is processed through MOPA **(3)** of the NDDPP **(1)**. FIG. 1 illustrates the input of PDS1 **(9)**, PDS2 **(10)**, MC **(13)**, and a DEM **(8)** instruction set into MOPA **(3)**. FIG. 3 illustrates input of PDS1 **(9)** and PDS2 **(10)** into MOPA-MM **(22)** and input of DEM **(8)** instructions set DEMM **(21)**, and MC **(13)**, if any, into MOPA-CM **(23)**. FIG. 6 Step **(31)** describes the input of data sets in MOPA **(3)**. PDS1 **(9)** and PDS2 **(10)** are input by MOPA-MM **(22)** and used to make matrix indices M1 **(11)** and M2 **(12)** as described in FIG. 6 Step **(32)** and as illustrated in FIG. 3. For the purposes of this example, M1-M1 will be an example of M1 **(11)**, and M2a-M1 and M2b-M1 are an example of M2 **(12)**. As described by FIG. 6 Step **(32)** and illustrated in FIG. 3, MOPA-MM **(22)** makes matrix indices M1-M1, M2a-M1 and M2b-M1 from the integer values in columns of PDS1-M1, PDS2a-M1, and PDS2b-M1, respectively. For every entry in this example, the value in each column of PDS1-M1 becomes the value of the corresponding index of M1-M1; the value in each column of PDS2a-M1 becomes the value of the corresponding index of M2a-M1; and the value in each column of PDS2b-M1 becomes the value of the corresponding index of M2b-M1.

In this example, the value of C1 in PDS1-M1 becomes the value of the corresponding first index of M1-M1, the value of C2 in PDS1-M1 becomes the value of the corresponding second index of M1-M1, etc. To illustrate, the first lines of PDS1-M1, PDS2a-M1, and PDS2b-M1 are:

```
2  4729    31    21     3     1     0     3     4
1     1    14     8     1     1     4     2     5
1   314    14    14     2     1     0     2     6
```

These example lines input to MOPA-MM **(22)** produce matrix indices:

```
M1-M1(  2  4729    31    21     3     1     0     3     4  )
M2a-M1( 1     1    14     8     1     1     4     2     5  )
M2b-M1( 1   314    14    14     2     1     0     2     6  )
```

All lines in the PDS1-M1, PDS2a-M1, and PDS2b-M1 are processed similarly in the MOPA-MM **(22)** to complete the three complete sets of matrix indices, M1-M1, M2a-M1 and M2b-M1. As illustrated in FIG. 3 and described in FIG 6 Steps **(32)** and **(33)**, MOPA-MM **(22)** outputs M1-M1, M2a-M1, and M2b-M1 to both USER **(5)** and MOPA-CM **(23)**. MOPA-CM **(23)** also inputs DEMM **(21)**. Since MC **(13)** is output data of NDDPP **(1)** and it is the first pass of data through NDDPP **(1)**, at this point in this example of the first embodiment there is no previously existing MC **(13)** input to the MOPA-CM **(23)**.

MOPA-CM **(23)** makes MC **(13)** as instructed by DEMM **(21)** and described in FIG. 6 Step **(33)**. DEMM **(21)** can instruct MOPA-CM **(23)** to output any combination of information related to shared indices of M1 **(11)** and M2 **(12)** and to use any previously produced MC **(13)** to further select information to be included in new MC **(13)** production. For the purposes of this example, DEMM-M1 is an example of DEMM **(21)** and MC-M1 will be an example of MC **(13)**. The example DEMM-M1 instructs MOPA-CM **(23)** to make MC-M1:

```
DEMM-M1:

IDENTIFY ALL MATRIX INDICES FROM M2-M
  (IN THIS EXAMPLE, M2a-M1 AND M2b-M1)
  THAT SHARE VALUES IN INDICES 7, 8 AND 9 WITH
  MATRIX INDICES 7, 8, AND 9 FROM M1-M
    (IN THIS EXAMPLE, M1-M1);AND
  PUT THESE INDICES IN MC-M
    (IN THIS EXAMPLE MC-M1)
```

MC-M1 matrix indices with the shared values 0, 3, and 4 are identified and put into MC-M1:

```
MC-M1( 2, 4729,   31,   21,    3,    1,    0,    3,     4  )
MC-M1( 1,  455,    8,    9,    2,   10,    0,    3,     4  )
MC-M1( 1,  456,   18,    9,    2,   10,    0,    3,     4  )
```

In this example of the first embodiment, index 2 of MC-M1 indicates the serial number of the atoms that share a five angstroms ("5A") on a side cubic volume whose corner position is indicated by indices 7, 8, and 9. In this example, the only matrix indices with values of indices 7, 8, and 9 that are shared between M1-M1 and the set of M2a-M1 and M2b-M1 have the values of 0, 3, 4, hereinafter referred to as "shared 5A volume 0, 3, 4". The atoms that occupy the shared 5A volume 0, 3, 4 are the atoms with the atom serial numbers 4729, 455, and 456. The indices with shared values are a spatial reference point for the volume, here a 5A-on-a-side cubic volume, that contains atoms from both input datasets, DS1-M1 and the set of DS2a-M1 and DS2b-M1.

MC **(13)** is output back to MOPA-CM **(23)** as shown in FIG. 3 and described in FIG. 6 Step **(34)**. As illustrated by FIG. 1 and FIG. 4 and described in FIG. 7 Step **(35)**, MC **(13)**, DS1 **(6)**, and DS2 **(7)** are input to FOPA **(4)**. MC **(13)** is used by FOPA-CC **(24)** to generate selection criteria SCF **(25)** as illustrated in FIG. 4 and described in FIG. 7 Step **(36)**. In this example, SCF-M1 is an example of SCF **(25)**:

```
SCF-M1:

SELECT ALL DS1-M1 AND DS2-M1 INPUT LINES THAT HAVE A
  C2 VALUE THAT MATCHES INDEX 2 IN MC-M1.
```

As illustrated in FIG. 4 and described in FIG. 7 Step **(37)**, FOPA-AC **(26)** uses SCF **(25)** to make DSC **(14)**. The following lines from Tables 1 and 3 match the selection criteria of SCF-M1 that are used by FOPA-AC **(26)** to make DSC-M1, an example of DSC **(14)**:

DSC-M1:

```
HETATM 4729  ZN  ZN • 1   0.004  -0.434  -5.873 1.00 0.00 ZN
ATOM    455  CE1 HIS  B 10   2.492   0.821  -7.457 1.00 0.00  C
ATOM    456  NE2 HIS  B 10   2.194  -0.032  -6.474 1.00 0.00  N
```

As illustrated in FIG. 4 and described in FIG. 7, Step **(38)**, DSC **(14)** is output to USER **(5)** and DIPA-PA **(17)**. FIG. 8 shows a view of the zinc atom **(39)** (atom serial number 4729), the NE2 atom **(40)** (atom serial number 456) and the CE1 atom **(41)** (atom serial number 455) in DSC-M1.

MC **(13)**, fed back to MOPA-CM **(23)**, can be used to further select information that goes into subsequent rounds of MC **(13)** production. For example, if MC **(13)** was fed back to MOPA-CM **(23)**, the information could be used to eliminate MC **(13)** entries being generated by MOPA-CM **(23)** that did not contain at least one zinc (matrix index 3 value 31), CE1 (matrix index 3 value 8), and NE2 (matrix index 3 value 18) atoms. In this way, for example, volumes with CE1, NE2, and zinc atoms can rapidly be identified as possible binding sites for zinc in subsequent processed protein structure data.

In a second example, 1AIY.PDB MODEL 1 insulin hexamer (chains A-L) is used as a new example of DS2 **(7)**, DS2-M2. For this example, 1AIY.PDB is stripped of all hydrogen atoms and any HETATMs, including water, other than ZN. DS2-M2 is processed as described above using DEMI-M1 and DEMS-M1, with the exception that -27.492 is substituted for C7LOW, -28.055 is substituted for C8LOW, and -38.947 is substituted for C9LOW. These values are specific to DS2-M2 and are substituted in the DEMS-M1 to make a new example of PDS2 **(10)**, PDS2-M2. In this secondary example, PDS2-M2 is processed by MOPA-MM **(22)** to make a new example of M2 **(12)**, M2-M2.

For this second example, a new DEMM **(21)**, DEMM-M2, is used to instruct the production of a new MC **(13)**.

DEMM-M2:

FOR ALL COMBINATIONS OF (a) AND (b) BELOW:

```
IF:
  a) A SET OF MATRIX INDICES FROM THE PREVIOUS MC-M
     (IN THIS EXAMPLE, MC-M1)
     SHARE ALL VALUES IN INDICES 7, 8, AND 9; AND
  b) A SET MATRIX INDICES FROM THE CURRENT M2-M
     (IN THIS EXAMPLE, M2-M2)
     SHARE ALL VALUES IN INDICES 7, 8, AND 9; AND
  c) THE VALUES OF INDEX 3 OF THE MATRIX INDICES IN (b)
     CONTAIN ALL OF THE VALUES FOUND IN
     INDEX 3 OF THE MATRIX INDICES IN (a)
     (IN THIS EXAMPLE, 8, 18, AND 31, THE INDEX 3 VALUES
      IN MC-M1)
THEN:
  PUT THE MATRIX INDICES IN (b) THAT HAVE A VALUE FOR
  INDEX 3 THAT MATCHES AN INDEX 3 VALUE IN MC-M1
  (IN THIS EXAMPLE, 8, 18, AND 31)
  INTO A NEW MC-M, MC-NEW; AND
WHEN ALL COMBINATIONS OF (a) AND (b) HAVE BEEN PROCESSED,
  WRITE ALL INDICES OF MC-NEW INTO MC-M; AND
  OUTPUT MC-M.
```

In this example, the example of MC **(13)**, MC-M2, is produced under instruction from DEMM-M2 using entries from the previous MC **(13)** and MC-NEW. MC-NEW contains matrix indices representing the atoms that occupy the shared 5A volume 5, 5, and 6:

```
MC-NEW(1,  455,    8,    9,    2,   10,    5,    5,     6  )
MC-NEW(1,  456,   18,    9,    2,   10,    5,    5,     6  )
MC-NEW(1, 2031,    8,    9,    7,   10,    5,    5,     6  )
MC-NEW(1, 2032,   18,    9,    7,   10,    5,    5,     6  )
MC-NEW(1, 3607,    8,    9,   11,   10,    5,    5,     6  )
MC-NEW(1, 3608,   18,    9,   11,   10,    5,    5,     6  )
MC-NEW(2, 4729,   31,   21,    3,    1,    5,    5,     6  )
```

MOPA-CM **(23)**, using DEMM **(21)**, puts matrix indices from MC-NEW into MC-

M2:

```
MC-M2( 1,  455,    8,    9,    2,   10,    5,    5,     6  )
MC-M2( 1,  456,   18,    9,    2,   10,    5,    5,     6  )
MC-M2( 1, 2031,    8,    9,    7,   10,    5,    5,     6  )
MC-M2( 1, 2032,   18,    9,    7,   10,    5,    5,     6  )
MC-M2( 1, 3607,    8,    9,   11,   10,    5,    5,     6  )
MC-M2( 1, 3608,   18,    9,   11,   10,    5,    5,     6  )
MC-M2( 2, 4729,   31,   21,    3,    1,    5,    5,     6  )
```

Using SCF-M1 as the example of SCF **(25)** in this example, then DSC **(14)**, in this example DSC-M2, is made:

```
ATOM    455 CE1 HIS  B 10   2.492   0.821  -7.457 1.00 0.00  C
ATOM    456 NE2 HIS  B 10   2.194  -0.032  -6.474 1.00 0.00  N
ATOM   2031 CE1 HIS  F 10  -2.113   1.204  -7.836 1.00 0.00  C
ATOM   2032 NE2 HIS  F 10  -1.304   1.308  -6.779 1.00 0.00  N
ATOM   3607 CE1 HIS  J 10  -0.106  -3.055  -7.971 1.00 0.00  C
ATOM   3608 NE2 HIS  J 10  -0.664  -2.435  -6.927 1.00 0.00  N
HETATM 4729  ZN  ZN  • 1   0.004  -0.434  -5.873 1.00 0.00 ZN
```

FIG. 9 shows a view of zinc atom **(39)** (atom serial number 4729), NE2 atom **(40)** (atom serial number 456), CE1 atom **(41)** (atom serial number 455), NE2 atom **(42)** (atom serial number 2032), CE1 atom **(43)** (atom serial number 2031), NE2 atom **(44)** (atom serial number 3608), and the CE1 atom **(45)** (atom serial number 3607) in DSC-M2 using C7, C8, and C9 as their cartesian coordinates. Shared occupancy of a volume by non-zinc atoms **(40)**, **(41)**, **(42)**, **(43)**, **(44)**, and **(45)** specified in DSC-M2 suggests that one example of a zinc binding site is comprised of three nitrogen atoms (NE2) presented to zinc by CE1 (carbon) atoms from three histidines in three different insulin chains.

FIG. 10 shows a view of a potential zinc binding site as defined by NE2 atom **(40)** (atom serial number 456), CE1 atom **(41)** (atom serial number 455), NE2 atom **(42)** (atom serial number 2032), CE1 atom **(43)** (atom serial number 2031), NE2 atom **(44)** (atom serial

number 3608), and the CE1 atom **(45)** (atom serial number 3607) in DSC-M2 using C7, C8, and C9 as their cartesian coordinates. Any volume populated by similar atoms from any input data set, DS1 **(6)** or DS2 **(7)**, would be a candidate for another zinc binding site independent of position or orientation.

Similarly, if an alternative DEMM **(21)** instructions set, DEMM-M2b, specified that "the matrix indices in (a) and (b) contain identical numbers for the value of index 4 but have different numbers for the value of matrix index 6", in addition to the instruction "the matrix indices in (a) and (b) contain at least three or more numbers that are the value of index 3 in common", then this alternative DEMM-M2b instruction set could be used to eliminate MC **(13)** entries being generated by MOPA-CM **(23)** that did not contain at least 2 atoms coming from two different histidines. Using an alternative DEMM-M2b, volumes with shared histidine atoms coming from more than one histidine could rapidly be identified as possible binding sites for zinc in subsequent processed protein structure data, even if no zinc were present. One skilled in the art will see that the DEMM **(21)** offers flexibility in specifying the matrix indices of interest and in controlling subsequent production of MC **(13)**. Atoms not found in the shared volume can be associated with the atoms in the shared volume by linking them in using the SCF **(25)**, in this case through their shared attribute of chain and residue designation. An example of this linking is illustrated using a specific alternative SCF **(25)** instruction set, SCF-M2b:

```
SCF-M2b:

a) SELECT ALL DS1-M AND DS2-M INPUT LINES THAT HAVE A
   C2 VALUE THAT MATCHES THE VALUE OF INDEX 2 IN MC-M1; AND
b) SELECT ALL DS1-M AND DS2-M INPUT LINES THAT HAVE
   THE SAME VALUE IN C5 AS THE VALUE OF THE
   MATRIX INDEX 5 IN (a) AND
   THE SAME VALUE IN C6 AS THE VALUE OF THE
   MATRIX INDEX 6 IN (a)
```

SCF-M2b would instruct FOPA-AC **(26)** to include all atoms in the residues that contribute atoms to the shared 5A volume 5, 5, 6 in DSC-M2b:

```
ATOM    447   N HIS B 10   6.449  -2.243  -7.785 1.00 0.00  N
ATOM    448  CA HIS B 10   6.423  -0.768  -7.975 1.00 0.00  C
ATOM    449   C HIS B 10   7.833  -0.273  -8.299 1.00 0.00  C
ATOM    450   O HIS B 10   8.018   0.820  -8.796 1.00 0.00  O
ATOM    451  CB HIS B 10   5.929  -0.100  -6.692 1.00 0.00  C
ATOM    452  CG HIS B 10   4.435   0.059  -6.753 1.00 0.00  C
ATOM    453 ND1 HIS B 10   3.819   0.907  -7.659 1.00 0.00  N
ATOM    454 CD2 HIS B 10   3.421  -0.516  -6.027 1.00 0.00  C
ATOM    455 CE1 HIS B 10   2.492   0.821  -7.457 1.00 0.00  C
ATOM    456 NE2 HIS B 10   2.194  -0.032  -6.474 1.00 0.00  N
ATOM   2023   N HIS F 10  -1.329   6.131  -7.960 1.00 0.00  N
ATOM   2024  CA HIS F 10  -2.614   5.426  -8.224 1.00 0.00  C
ATOM   2025   C HIS F 10  -3.704   6.451  -8.546 1.00 0.00  C
ATOM   2026   O HIS F 10  -4.743   6.116  -9.079 1.00 0.00  O
ATOM   2027  CB HIS F 10  -3.022   4.625  -6.986 1.00 0.00  C
ATOM   2028  CG HIS F 10  -2.439   3.241  -7.063 1.00 0.00  C
ATOM   2029 ND1 HIS F 10  -2.811   2.334  -8.044 1.00 0.00  N
ATOM   2030 CD2 HIS F 10  -1.508   2.594  -6.288 1.00 0.00  C
ATOM   2031 CE1 HIS F 10  -2.113   1.204  -7.836 1.00 0.00  C
ATOM   2032 NE2 HIS F 10  -1.304   1.308  -6.779 1.00 0.00  N
ATOM   3599   N HIS J 10  -4.683  -4.963  -8.227 1.00 0.00  N
ATOM   3600  CA HIS J 10  -3.420  -5.713  -8.478 1.00 0.00  C
ATOM   3601   C HIS J 10  -3.748  -7.170  -8.813 1.00 0.00  C
ATOM   3602   O HIS J 10  -2.923  -7.898  -9.329 1.00 0.00  O
ATOM   3603  CB HIS J 10  -2.536  -5.662  -7.228 1.00 0.00  C
ATOM   3604  CG HIS J 10  -1.680  -4.426  -7.266 1.00 0.00  C
ATOM   3605 ND1 HIS J 10  -0.682  -4.248  -8.209 1.00 0.00  N
ATOM   3606 CD2 HIS J 10  -1.659  -3.299  -6.480 1.00 0.00  C
ATOM   3607 CE1 HIS J 10  -0.106  -3.055  -7.971 1.00 0.00  C
ATOM   3608 NE2 HIS J 10  -0.664  -2.435  -6.927 1.00 0.00  N
HETATM 4729  ZN  ZN • 1   0.004  -0.434  -5.873 1.00 0.00 ZN
```

FIG. 11 shows a view of three histidine residues, listed in DSC-M2b, that contribute CE1 and ND2 atoms to a shared volume with atom 4729, zinc **(39)**: histidine residue number 10 from the B chain **(46)**, histidine residue number 10 from the F chain **(47)**, and histidine residue number 10 from the J chain **(48)**.

In a third example, 1AIY.PDB MODEL 1 insulin hexamer (chains A-L) is processed as described above using DEMI-M1 and DEMS-M1, with the exception that -27.024 is C7LOW, -24.983 is C8LOW, and -38.947 is C9LOW. When the entire insulin hexamer from 1AIY.PDB is analyzed as outlined above, the shared occupancy of the atoms of a bound molecule can be linked and all residues in a protein that share a volume with that bound molecule can be readily found and output as DSC-M3, an example of DSC **(14)**.

In DSC-M3, phenol **(49)** and insulin hexamer atoms from shared 5A volume 7, 5, 7 are:

```
ATOM     79   C CYS A  6  12.718   0.513  -1.289 1.00 0.00  C
ATOM     80   O CYS A  6  11.737   1.189  -1.528 1.00 0.00  O
ATOM    113   C SER A  9  12.033   4.250   0.443 1.00 0.00  C
ATOM    114   O SER A  9  12.954   3.780  -0.194 1.00 0.00  O
ATOM    131  HA ILE A 10  11.289   4.893  -2.151 1.00 0.00  H
ATOM   1917 CG1 VAL F  2   8.772   2.869   0.866 1.00 0.00  C
ATOM   1924 3HG1 VAL F  2   9.260   2.768  -0.091 1.00 0.00  H
HETATM 4737  O1 IPH •  3  11.886   3.494  -3.679 1.00 0.00  O
HETATM 4738  H2 IPH •  3   9.331   2.845  -3.504 1.00 0.00  H
HETATM 4743 HO1 IPH •  3  11.634   2.817  -3.047 1.00 0.00  H
```

In DSC-M3, phenol **(49)** and insulin hexamer atoms from shared 5A volume 7, 5, 6 are:

```
ATOM    450   O HIS B 10   8.018   0.820  -8.796 1.00 0.00  O
ATOM    471 CD2 LEU B 11  10.326   0.382  -5.618 1.00 0.00  C
ATOM    473  HA LEU B 11  10.289   0.439  -8.272 1.00 0.00  H
ATOM    480 1HD2 LEU B 11  11.172   0.981  -5.921 1.00 0.00  H
```

```
ATOM    4812HD2 LEU  B 11   9.452   0.683  -6.177 1.00 0.00  H
ATOM    4823HD2 LEU  B 11  10.143   0.525  -4.563 1.00 0.00  H
HETATM 4731  C1 IPH  • 3  10.995   3.444  -4.732 1.00 0.00  C
HETATM 4732  C2 IPH  • 3   9.662   3.084  -4.504 1.00 0.00  C
HETATM 4733  C3 IPH  • 3   8.758   3.034  -5.572 1.00 0.00  C
HETATM 4734  C4 IPH  • 3   9.187   3.344  -6.868 1.00 0.00  C
HETATM 4735  C5 IPH  • 3  10.520   3.704  -7.097 1.00 0.00  C
HETATM 4736  C6 IPH  • 3  11.425   3.754  -6.029 1.00 0.00  C
HETATM 4740  H4 IPH  • 3   8.490   3.306  -7.691 1.00 0.00  H
HETATM 4741  H5 IPH  • 3  10.851   3.943  -8.096 1.00 0.00  H
HETATM 4742  H6 IPH  • 3  12.453   4.033  -6.205 1.00 0.00  H
```

In DSC-M3, phenol **(49)** and insulin hexamer atoms from shared 5A volume 6, 5, 6 are:

```
ATOM    452  CG HIS  B 10   4.435   0.059  -6.753 1.00 0.00  C
ATOM    453 ND1 HIS  B 10   3.819   0.907  -7.659 1.00 0.00  N
ATOM    460 2HB HIS  B 10   6.391   0.872  -6.591 1.00 0.00  H
ATOM    461 HD1 HIS  B 10   4.269   1.470  -8.325 1.00 0.00  H
ATOM   1983 CD2 LEU  F  6   5.188   3.266  -4.546 1.00 0.00  C
ATOM   1985  HA LEU  F  6   2.989   4.806  -5.545 1.00 0.00  H
ATOM   1992 1HD2 LEU  F  6   4.810   2.591  -5.300 1.00 0.00  H
ATOM   1993 2HD2 LEU  F  6   5.268   4.261  -4.961 1.00 0.00  H
ATOM   1994 3HD2 LEU  F  6   6.161   2.933  -4.220 1.00 0.00  H
HETATM 4739  H3 IPH  • 3   7.729   2.756  -5.396 1.00 0.00  H
```

The residues that make up the binding pocket for the phenol molecule are assembled into DSC **(14)** output, where the DEMM **(21)** instruction set returns all DS1 **(6)** and DS2 **(7)** residue atoms for any residue containing an atom in the volume shared with phenol **(49)** atoms. This example of the result of volume linking through shared molecule occupancy is shown in FIG. 12. FIG. 12 phenol molecules **(49)**, **(50)**, **(51)** from the 1AIY insulin hexamer model 1 are surrounded by residues from the insulin hexamer: the A chains (**(64)**, **(65)**, **(66)**), the B chain (**(52)**, **(53), (54), (55)**), the E chain (**(67)**, **(68)**, **(69)**), the F chain (**(56)**, **(57), (58),**

(59)), the I chain (**(70)**, **(71)**, **(72)**), and the J chain (**(60)**, **(61), (62), (63)**).

In FIG. 12, the phenol molecule **(49)** contacts residues from:

chain A: CYS 6 **(64)**, SER 9 **(65)**, and ILE 10 **(66)**;

chain B: HIS 10 **(54)** and LEU 11 **(55)**; and

chain F: VAL 2 **(56)** LEU 6 **(57)**.

In FIG. 12, the phenol molecule **(50)** contacts residues from:

chain E: CYS 6 **(67)**, SER 9 **(68)**, and ILE 10 **(69)**; and

chain F: HIS 10 **(58)** and LEU 11 **(59)**; and

chain J: VAL 2 **(60)** LEU 6 **(61)**.

In FIG. 12 , the phenol molecule **(51)** contacts residues from:

chain I: CYS 6 **(70)**, SER 9 **(71)**, and ILE 10 **(72)**;

chain J: HIS 10 **(62)** and LEU 11 **(63)**;

chain B: VAL 2 **(52)** LEU 6 **(53)**.

As can be seen from the description of the examples of the first embodiment, the effect of throwing a volume grid consisting of cubes that are five angstroms on a side over the molecule in positive space is sufficient to return volumes that contain multiple atoms from the input sets. In addition, a catalogue of occupancy and contacts can be made from matrix indices representing common spatial occupancy of n-dimensional data elements. Once the atoms that contribute to common spatial occupancy are found, it is possible to search for similar atom clusters in other volumes in the same data set or in other data sets. This means that a search for binding site configurations can be done without resorting to calculations of distances between atoms or surface points. The complete environments of bound molecules can be constructed from linking sets of atoms with shared occupancy of volumes through their common contact with a bound molecule. This is accomplished through use of the MC **(13)** to select input data.

In order to illustrate an example of DSC **(14)** feedback to DIPA **(2)**, DSC-M1 output from the analysis of DS1-M1, DS2a-M1, and DS2b-M1 will be used as an example of DSC **(14)** that is fed back to DIPA-PA **(17)** as input and referred to as DS1-M4. A new PDB file, 3D1M.PDB, will be used as DS2 **(7)** input, in this example and referred to as DS2-M4. 3D1M.PDB is a data set containing structural and chemical information on a signaling/cell

adhesion protein sonic hedgehog. For this example, 3D1M.PDB is stripped of all hydrogen atoms and any HETATMs, including water, other than ZN. DS1-M4 and DS2-M4 are, in this example, processed separately by DIPA **(2).** DS2-M4 is processed as described for DS1-M1 but uses different SHIFT values (0.000 substituted for C7LOW, -60.000 substituted for C8LOW and -8.309 substituted for C9LOW) and different MULTFACTOR values (0.1 substituted for 0.2). MOPA-MM **(22)** creates M2-M4, an example of M2 **(12)**. MOPA-CM **(23)** creates MC **(13)** using DEMM **(21)**. In this example, a specific DEMM-M4, instructs MOPA-CM **(23)** to make MC **(13)**, in this example, MC-M4.

```
DEMM-M4:

FOR ALL COMBINATIONS OF (a) AND (b) BELOW:
IF:
  a) A SET OF MATRIX INDICES FROM MC-M1
     SHARE ALL VALUES IN INDICES 7, 8, AND 9; AND
  b) A SET OF MATRIX INDICES FROM THE M2-M4
     SHARE ALL VALUES IN INDICES 7, 8, AND 9; AND
  c) THE VALUES OF INDEX 3 OF THE MATRIX INDICES IN (b)
     CONTAIN ALL OF THE VALUES FOUND IN INDEX 3 OF THE
     MATRIX INDICES IN (a);
THEN:
  PUT THE MATRIX INDICES IN (b) THAT HAVE A VALUE FOR
  INDEX 3 THAT IS ONE OF: 8, 15, 18, AND 31
  (THE INDEX 3 VALUES IN MC-M1 AND VALUE 15,
   AN EQUIVALENT VALUE TO VALUE 18 IN THIS EXAMPLE)
  INTO A NEW MC (13), MC-NEW; AND
WHEN ALL COMBINATIONS OF (a) AND (b) HAVE BEEN PROCESSED,
  WRITE ALL INDICES OF MC-NEW INTO MC-M; AND
  OUTPUT MC-M.
```

In this specific example of using DCS **(14)** input to DIPA **(2)**, DSC-M1 contains atoms that have been determined to have common spatial occupancy of shared 5A volume 0, 3, 4. MC-M4 contains matrix indices related to atoms that have been determined to have a

common spatial occupancy of shared 10A volume 4, 7, 2 and shared 10A volume 5, 2, 2. There is no intrinsic relationship between the reference volumes containing M1-M1 5A shared volume (0, 3, 4) and either of the MC-M4 10A shared volumes ((4, 7, 2) and (5, 2, 2)). Once M1 **(11)** and M2 **(12)** are made in MOPA-MM **(22)** then DEMM-M4 directs MOPA-CM **(23)** to make MC **(13)**. M1-M4, the example of M1**(11)** in this example, contains matrix indices, (a), that are based on atoms that occupy shared volume 0, 3, 4. M2-M4, the example of M2 **(12)** in this example, contains matrix indices, (b), that are based on atoms that occupy two different shared volumes, shared volume 4, 7, 2 and shared volume 5, 2, 2. An example of MC **(13)**, MC-M4, is made that represents the common spatial occupancy of certain atoms in different sized volumes in different reference frames. DEMM-M4 instructs MOPA-CM to output the matrix indices from M2-M4 that have shared values of indices 7, 8, and 9 and have the same or equivalent values in index three as the atoms that are associated with shared volume 0, 3, 4:

```
MC-M4:

MC-M4( 1, 1959,   15,    9,    2,  135,    4,    7,     2  )
MC-M4( 1, 1961,    8,    9,    2,  135,    4,    7,     2  )
MC-M4( 1, 2007,   15,    9,    2,  141,    4,    7,     2  )
MC-M4( 1, 2009,    8,    9,    2,  141,    4,    7,     2  )
MC-M4( 1, 2010,   18,    9,    2,  141,    4,    7,     2  )
MC-M4( 1, 2335,    8,    9,    2,  181,    4,    7,     2  )
MC-M4( 1, 2336,   18,    9,    2,  181,    4,    7,     2  )
MC-M4( 1, 2351,   15,    9,    2,  183,    4,    7,     2  )
MC-M4( 1, 2353,    8,    9,    2,  183,    4,    7,     2  )
MC-M4( 1, 2354,   18,    9,    2,  183,    4,    7,     2  )
MC-M4( 2, 3919,   31,   21,    2,    2,    4,    7,     2  )

MC-M4( 1,  763,    8,    9,    1,  135,    5,    2,     2  )
MC-M4( 1,  809,   15,    9,    1,  141,    5,    2,     2  )
MC-M4( 1,  811,    8,    9,    1,  141,    5,    2,     2  )
MC-M4( 1,  812,   18,    9,    1,  141,    5,    2,     2  )
```

```
MC-M4( 1, 1153,   15,    9,    1,  183,    5,    2,     2  )
MC-M4( 1, 1155,    8,    9,    1,  183,    5,    2,     2  )
MC-M4( 1, 1156,   18,    9,    1,  183,    5,    2,     2  )
MC-M4( 2, 3916,   31,   21,    1,    1,    5,    2,     2  )
```

MC-M4 is then used to make a specific example SCF **(25)**, SCF-M4, to select lines from 3D1M.PDB that are output as DSC-M4, an example of DSC **(14)**.

```
SCF-M4:

SELECT ALL DS2-M4 INPUT LINES THAT HAVE A C3 VALUE
OF 8 (CE1), 15 (ND1), 18 (NE2) OR 31 (ZN).
```

After completion of processing by NDDPP **(1)**, DSC-M4 is output:

```
ATOM   1959 ND1 HIS B135 48.363  11.262  14.489 1.00 28.38  N
ATOM   1961 CE1 HIS B135 49.494  11.838  14.849 1.00 30.74  C
ATOM   2007 ND1 HIS B141 49.187  15.587  13.482 1.00 30.13  N
ATOM   2009 CE1 HIS B141 48.737  15.750  14.716 1.00 27.66  C
ATOM   2010 NE2 HIS B141 47.684  14.969  14.882 1.00 28.38  N
ATOM   2335 CE1 HIS B181 47.604  12.494  21.584 1.00 32.28  C
ATOM   2336 NE2 HIS B181 47.177  13.509  20.844 1.00 32.03  N
ATOM   2351 ND1 HIS B183 44.840  15.673  16.395 1.00 31.00  N
ATOM   2353 CE1 HIS B183 44.346  15.809  15.173 1.00 30.53  C
ATOM   2354 NE2 HIS B183 43.303  16.621  15.232 1.00 28.51  N
HETATM 3919  ZN  ZN  B  2  46.654  14.418  16.674 0.50 26.39 ZN

ATOM    763 CE1 HIS A135 50.424 -30.421  16.372 1.00 33.62  C
ATOM    809 ND1 HIS A141 50.856 -34.602  15.542 1.00 34.11  N
ATOM    811 CE1 HIS A141 51.212 -34.651  16.816 1.00 34.72  C
ATOM    812 NE2 HIS A141 52.195 -33.788  17.011 1.00 33.96  N
ATOM   1153 ND1 HIS A183 54.859 -34.825  18.660 1.00 31.25  N
ATOM   1155 CE1 HIS A183 55.424 -35.007  17.480 1.00 31.21  C
```

```
ATOM   1156  NE2 HIS  A183 56.454 -35.822  17.633  1.00 30.37   N
HETATM 3916   ZN   ZN  A  1  53.235 -33.247  18.916  0.50 25.53  ZN
```

As an example, atoms from the shared 10A volume 4, 7, 2 with the following atom serial numbers listed are associated with the corresponding listed points shown in FIG. 13 as follows:

Atom Serial Number 3919: ZN **(73)**

Atom Serial Number 2010: NE2 **(74)**

Atom Serial Number 2009: CE1 **(75)**

Atom Serial Number 2351: ND1 **(76)**

Atom Serial Number 2353: CE1 **(77)**

Atom Serial Number 1959: ND1 **(78)**

Atom Serial Number 1961: CE1 **(79)**

Atom Serial Number 2336: NE2 **(80)**

Atom Serial Number 2335: CE1 **(81)**

The distance **(82)** between the center of the zinc atom **(73)** and the nitrogen ND1 **(78)** is 4.202 A. The distance **(83)** between the center of the zinc atom **(73)** and the nitrogen NE2 **(80)** is 4.300 A. Distances **(82)** and **(83)** are much larger than the distances between **(73)** and **(74)**, 2.139 A, and **(73)** and **(76)**, 2.223 A. As can be seen from FIG. 13, the nitrogen atoms identified in DSC-M4, (**74**), **(76)**, **(78)**, and **(80)**, are oriented less symmetrically than those in FIG. 10 nitrogen atoms **(40)**, **(42)**, and **(44)**. CE1 atoms shown in FIG. 13 **(75)**, **(77**), and **(79)** still present bound ND1 and NE2 atoms **(74)**, **(76)**, **(78)**, and **(80)** to interact with the zinc atom **(73)**. What is interesting about the presentation of ND1 and NE2 atoms in FIG. 13 is that there are two nitrogen atoms, **(74)** and **(76)**, in direct contact (within 2.3A) with the zinc **(73)**, and two nitrogen atoms **(78)** and **(80)** that are close (within 4.3 A) but not in direct contact with the zinc **(73)**. Both close nitrogen atoms **(78)** and **(80)** are roughly equidistant to the zinc atom **(73)** and either could move closer to coordinate the ZN with two nitrogens **(74)** and **(76)** in direct contact. This may form the basis for a histidine-zinc "switch" with coordination between one set of three histidines **(74)**, **(76)**, and **(80)** stabilizing one state and another set of three histidines stabilizing a different state.

Using a different MULTFACTOR value, 0.1, for the processing of DS1-M4, than for DS1-M1, 0.2, essentially throws a wider "net" for structures in input DS1-M4, 3D1M.PDB, for atoms that were found to occupy a smaller common spatial volume in DSC-M1. The particular instructions used in this example of NDDPP **(1)** processing help to pinpoint what atoms coordinate zinc atom binding in proteins but more importantly help to identify possible roles that zinc binding might have with respect to the function of the protein.

1MEY.PDB is a data set containing structural and chemical information on a zinc finger transferase protein. For this example, 1MEY.PDB is stripped of all hydrogen atoms and any HETATMs, including water, other than ZN. 1MEY.PDB is input as an example of DS2 **(7)**, DS2-M5, and processed using DEMI-M1 and DEMS-M1 with the exception that -1.605 is substituted for C7LOW, 0.000 is substituted for C8LOW, and -9.277 is substituted for C9LOW. MOPA-CM **(23)** creates MC **(13)** using DEMM **(21)**. In this example, DEMM-M5 is an example of DEMM **(21)** and instructs MOPA-CM **(23)** to make MC **(13)**, in this example, MC-M5:

```
DEMM-M5:

FOR ALL COMBINATIONS OF (a) AND (b) BELOW:
IF:
  a) A SET OF MATRIX INDICES FROM MC-M1
     SHARE ALL VALUES IN INDICES 7, 8, AND 9; AND
  b) A SET MATRIX INDICES FROM M2-M5
     SHARE ALL VALUES IN INDICES 7, 8, AND 9; AND
  c) THE VALUES OF INDEX 3 OF THE MATRIX INDICES IN (b)
     CONTAIN ALL OF THE VALUES FOUND IN INDEX 3 OF THE
     MATRIX INDICES IN (a);
THEN:
  PUT THE MATRIX INDICES IN (b) INTO A NEW MC-M,
  MC-NEW; AND
WHEN ALL COMBINATIONS OF (a) AND (b) HAVE BEEN PROCESSED,
  WRITE ALL INDICES OF MC-NEW INTO MC-M; AND
  OUTPUT MC-M.
```

MC-M5 is made using DEMM-M5:

```
MC-M5( 1, 1792,    3,    5,    7,    7,   10,   18,     3  )
MC-M5( 1, 1793,   30,    5,    7,    7,   10,   18,     3  )
MC-M5( 1, 1815,   30,    5,    7,   10,   10,   18,     3  )
MC-M5( 1, 1915,    8,    9,    7,   23,   10,   18,     3  )
MC-M5( 1, 2032,   18,    9,    7,   23,   10,   18,     3  )
MC-M5( 1, 3607,    8,    9,    7,   27,   10,   18,     3  )
MC-M5( 1, 3608,   18,    9,    7,   27,   10,   18,     3  )
MC-M5( 2, 4729,   31,   21,    7,   88,   10,   18,     3  )
```

MC-M5 is then used to make a specific example SCF **(25)**, SCF-M5, to select lines from 3D1M.PDB that are output as DSC-M5.

```
SCF-M5:

SELECT ALL DS2-M4 INPUT LINES THAT HAVE A C2 VALUE
THAT IS THE SAME AS THE VALUE OF INDEX 2 OF MC-M5
```

DSC-M5 is output using SCF-M5:

```
ATOM   1792  CB CYS  F  7  50.576  94.574   6.895  1.00 50.85  C
ATOM   1793  SG CYS  F  7  51.487  93.991   8.367  1.00 47.79  S
ATOM   1815  SG CYS  F 10  51.452  90.484   6.506  1.00 45.98  S
ATOM   1915 CE1 HIS  F 23  48.396  90.153   8.450  1.00 37.97  C
ATOM   1916 NE2 HIS  F 23  49.066  91.263   8.674  1.00 43.30  N
ATOM   1952 CE1 HIS  F 27  53.111  90.230  10.338  1.00 56.95  C
ATOM   1953 NE2 HIS  F 27  52.037  90.966  10.110  1.00 52.37  N
HETATM 2643  ZN  ZN  F 88  51.025  91.575   8.451  1.00 55.01 ZN
```

Atoms with the following atom serial numbers listed are associated with the corresponding listed points shown in FIG. 14 as follows:

Atom Serial Number 2643: ZN **(84)**

Atom Serial Number 1916: NE2 **(85)**

Atom Serial Number 1915: CE1 **(86)**

Atom Serial Number 1953: NE2 **(87)**

Atom Serial Number 1952: CE1 **(88)**

Atom Serial Number 1793: SG **(89)**

Atom Serial Number 1815: SG **(90)**

The distance between the center of the zinc atom **(84)** and:

the nitrogen NE2 **(85)** is 1.996 A,

the nitrogen NE2 **(87)** is 2.036 A,

the sulfur SG **(89)** is 2.461 A, and

the sulfur SG **(90)** is 2.271 A.

As can be seen by an examination of FIG. 14, two SG atoms, **(89)** and **(90)**, from two cysteines and two NE2 atoms, **(85)** and **(87)**, from two histidines in 1MEY.PDB are bound in a tetrahedral configuration to a zinc atom **(84)**. As can be seen from FIG. 14 and the distances above, the distances between the bound zinc and the coordinating atoms **(85)**, **(87)**, **(89)** and **(90)** are roughly equal. The presence of sulfur atoms, instead of NE2 or ND1 atoms as in 1AIY.PDB and 3D1M.PDB, may provide a structural advantage for proteins exposed to molecules that provide strong electrostatic interactions such as DNA. DSC-M5 demonstrates the repeated use of NDDPP **(1)** to understand what components are present in shared volumes.

In order to demonstrate the use of systematically changing values in DEMS **(16)**, an example of incremented SHIFT values will be described. In this example, 3M7K.PDB is input as an example of DS1 **(6)**, DS1-M6. For this example, 3M7K.PDB is stripped of all hydrogen atoms and any HETATMs, including water, other than ZN. 3M7K.PDB is a data set containing structural and chemical information on a restriction endonuclease (hydrolase), PAC1. The example of DEMI **(15)** in this example is DEMI-M6:

```
DEMI-M6:
```

```
FOR ALL DS1-M ENTRIES:

DATA PARSING:
  +C1 +C2 +C3 +C4 +C5 +C6 +C7 +C8 +C9
  -C10 -C11 -C12

DATA SEPARATION:
  TABS AND LINE RETURNS
  C1 THROUGH C12:
  1 RECORD NAME (“ATOM” or “HETATM”)
  2 ATOM SERIAL NUMBER IN THE DATA
  3 ATOM NAME
  4 RESIDUE NAME
  5 CHAIN IDENTIFIER
  6 RESIDUE SEQUENCE NUMBER
  7 ORTHOGONAL COORDINATES - atoms's x position (angstroms)
  8 ORTHOGONAL COORDINATES - atoms's y position (angstroms)
  9 ORTHOGONAL COORDINATES - atoms's x position (angstroms)
  10 OCCUPANCY
  11 TEMPERATURE FACTOR
  12 ELEMENT SYMBOL

FOR C7, C8, AND C9
  CALCULATE C7LOW, C8LOW, AND C9LOW
```

After parsing and attributing by DEMI **(15)**, the following C7LOW, C8LOW and C9LOW values are obtained by the execution of DEMI **(15)** in DIPA-PA **(17)**:

C7LOW = - 0.833

C8LOW = 13.098

C9LOW = 19.665

In this example, the example DEMI-M6 further instructs the DIPA-PA **(17)**:

```
DEMI-M6 (continued)
```

```
FOR ALL PADS1-M1 DATA ENTRIES:

GENERATE PDS1-M6-1 USING SHIFTADD = 0
GENERATE PDS1-M6-2 USING SHIFTADD = 1
GENERATE PDS1-M6-3 USING SHIFTADD = 3

C1 TO C1, C2 TO C2, C3 TO C3, C4 TO C4, C5 TO C5, C6 TO C6

MULTFACTOR = 0.2

SHIFT COMMANDS:
  IF THE SMALLEST VALUE OF C7(“C7LOW”)IS NEGATIVE, THEN
    SUBTRACT THE VALUE OF C7LOW FROM EVERY C7 ENTRY AND
    PUT THE RESULT IN C7
  ELSE
    PUT C7 IN C7
  ADD SHIFTADD TO C7 AND PUT RESULT IN C7
  IF THE SMALLEST VALUE OF C8(“C8LOW”)IS NEGATIVE, THEN
    SUBTRACT THE VALUE OF C8LOW FROM EVERY C8 ENTRY AND
    PUT THE RESULT IN C8
  ELSE
    PUT C8 IN C8
  ADD SHIFTADD TO C8 AND PUT RESULT IN C8
  IF THE SMALLEST VALUE OF C9(“C9LOW”)IS NEGATIVE, THEN
    SUBTRACT THE VALUE OF C9LOW FROM EVERY C9 ENTRY AND
    PUT THE RESULT IN C9
  ELSE
    PUT C9 IN C9

SIZE COMMANDS:
  MULTIPLY THE CURRENT VALUE OF C7 BY MULTFACTOR; AND
    PUT THE RESULT IN C7
  MULTIPLY THE CURRENT VALUE OF C8 BY MULTFACTOR; AND
    PUT THE RESULT IN C8
```

```
    MULTIPLY THE CURRENT VALUE OF C9 BY MULTFACTOR; AND
      PUT THE RESULT IN C9

  TRUNCATE COMMANDS:
    TRUNCATE THE CURRENT VALUE OF C7 AFTER THE DECIMAL PLACE;
      AND PUT THE INTEGER RESULT IN C7
    TRUNCATE THE CURRENT VALUE OF C8 AFTER THE DECIMAL PLACE;
      AND PUT THE INTEGER RESULT IN C8
    TRUNCATE THE CURRENT VALUE OF C9 AFTER THE DECIMAL PLACE;
      AND PUT THE INTEGER RESULT IN C9
```

In this example, for the first PDS1 **(9)** produced, PDS1-M6-1, the SHIFT VALUE used for C7 is -0.833, for C8 is 0.000 and for C9 is 0.000. For the second PDS1 **(9)** produced, PDS1-M6-2, the SHIFT VALUE used for C7 is -1.833, for C8 is -1.000 and for C9 is 0.000. For the second PDS1 **(9)** produced, PDS1-M6-2, the SHIFT VALUE used for C7 is -3.833, for C8 is -3.000 and for C9 is 0.000. In this example, MOPA-MM **(22)** makes M1-M6-1 from PDS1-M6-1, M1-M6-2 from PDS1-M6-2, and M1-M6-3 from PDS1-M6-3.

In this example, MC-M5 is input to MOPA-CM **(23)** as an example of MC **(13)** being fed back to MOPA-CM **(23)**. M1-M6-1, M1-M6-2, and M1-M6-3 are each processed separately as examples of M1 **(11)**. DEMM-M6, a specific example of DEMM **(21)**, instructs MOPA-CM **(23)** to make MC-M6-1, MC-M6-2, and MC-M6-3:

```
  DEMM-M6:

  FOR ALL COMBINATIONS OF (a) AND (b) BELOW:
  IF:
    a) A SET OF MATRIX INDICES FROM MC-M5
       SHARE ALL VALUES IN INDICES 7, 8, AND 9; AND
    b) A SET MATRIX INDICES FROM THE M1-M
       SHARE ALL VALUES IN INDICES 7, 8, AND 9; AND
    c) AT LEAST SIX OF THE VALUES OF INDEX 3 OF THE MATRIX
       INDICES IN (b) ARE ONE OF THE VALUES FOUND IN INDEX 3
       OF THE MATRIX INDICES IN (a) AND INCLUDE AT LEAST ONE
```

```
      VALUE OF INDEX 3 THAT IS 31;
  THEN:
    PUT THE MATRIX INDICES IN (b)
    INTO A NEW MC-M, MC-NEW; AND
  WHEN ALL COMBINATIONS OF (a) AND (b) HAVE BEEN PROCESSED,
    WRITE ALL INDICES OF MC-NEW INTO MC (13); AND
    OUTPUT MC-M.
```

MC-M6-1 is empty. MC-M6-2 contains:

```
MC-M6-2(1, 509,   30,    5,    1,   63,    4,    5,     7  )
MC-M6-2(1, 531,   30,    5,    1,   66,    4,    5,     7  )
MC-M6-2(1, 873,   30,    5,    1,  109,    4,    5,     7  )
MC-M6-2(1, 894,    2,    5,    1,  112,    4,    5,     7  )
MC-M6-2(1, 897,    3,    5,    1,  112,    4,    5,     7  )
MC-M6-2(1, 898,   30,    5,    1,  112,    4,    5,     7  )
MC-M6-2(2,1522,   31,   21,    1,  143,    4,    5,     7  )
```

MC-M6-1 is empty. MC-M6-3 contains:

```
MC-M6-2(1,  29,    3,    5,    1,    4,    5,    8,    11  )
MC-M6-2(1,  30,   30,    5,    1,    4,    5,    8,    11  )
MC-M6-2(1,  54,   30,    5,    1,    7,    5,    8,    11  )
MC-M6-2(1, 188,    3,    5,    1,   24,    5,    8,    11  )
MC-M6-2(1, 189,   30,    5,    1,   24,    5,    8,    11  )
MC-M6-2(1, 207,    3,    5,    1,   27,    5,    8,    11  )
MC-M6-2(1, 208,   30,    5,    1,   27,    5,    8,    11  )
MC-M6-2(2,1523,   31,   21,    1,  144,    5,    8,    11  )
```

MC-M6-2 is then used to make a specific example SCF **(25)**, SCF-M6-2, to select lines from 3M7K.PDB that are output as DSC-M6-2:

```
SCF-M6-2:
```

```
SELECT ALL DS2-M6-2 INPUT LINES THAT HAVE A C2 VALUE
THAT IS THE SAME AS THE VALUE OF INDEX 2 OF MC-M6
```

DSC-M6-2 is output using SCF-M6-2:

```
ATOM    509  SG CYS A 63  19.932  25.930  39.673  1.00 32.33  S
ATOM    531  SG CYS A 66  22.782  24.478  37.826  1.00 36.24  S
ATOM    873  SG CYS A109  22.510  28.378  37.833  1.00 27.98  S
ATOM    894  CA CYS A112  19.104  28.769  35.223  1.00 25.19  C
ATOM    897  CB CYS A112  18.961  27.620  36.231  1.00 25.41  C
ATOM    898  SG CYS A112  20.089  26.259  35.887  1.00 31.08  S
HETATM 1522  ZN  ZN A143  21.420  26.313  37.818  1.00 23.82 ZN
```

MC-M6-3 is then used to make a specific example SCF **(25)**, SCF-M6-3, to select lines from 3M7K.PDB that are output as DSC-M6-3:

```
SCF-M6-3:

SELECT ALL DS1-M6-3 INPUT LINES THAT HAVE A C2 VALUE
THAT IS THE SAME AS THE VALUE OF INDEX 2 OF MC-M6
```

DSC-M6-3 is output using SCF-M6-3:

```
ATOM     29  CB CYS A  4  25.587  39.556  59.417  1.00 56.18  C
ATOM     30  SG CYS A  4  24.733  40.661  58.213  1.00 53.91  S
ATOM     54  SG CYS A  7  22.350  38.237  59.719  1.00 51.46  S
ATOM    188  CB CYS A 24  25.486  37.888  55.771  1.00 43.60  C
ATOM    189  SG CYS A 24  23.712  37.702  56.055  1.00 42.94  S
ATOM    207  CB CYS A 27  21.468  40.760  55.083  1.00 39.57  C
ATOM    208  SG CYS A 27  21.173  40.388  56.852  1.00 41.95  S
HETATM 1523  ZN  ZN A144  23.033  39.268  57.604  1.00 37.76 ZN
```

The atoms in DSC-M6-2 do not show up in DSC-M6-3 because the DSC-M6-2 atoms are split between two volumes and do not satisfy the DEMM-M6 selection criteria. Similarly, the atoms in DSC-M6-3 do not show up in DSC-M6-2 because the DSC-M6-3 atoms are split between two volumes and do not satisfy the DEMM-M6 selection criteria. By incrementing the shift values, both zinc centers in the protein, 3M7K.PDB, and their associated SG atoms are found. Figure 15 shows a view of a graphical display of these DSC-M6-2 and DSC6-3 lines. Atoms with the following atom serial numbers listed are associated with the corresponding listed points shown in FIG. 14 as follows:

Atom Serial Number 1522: ZN **(91)**
Atom Serial Number 509: SG **(92)**
Atom Serial Number 873: SG **(93)**
Atom Serial Number 898: SG **(94)**
Atom Serial Number 531: SG **(95)**
Atom Serial Number 1523: ZN **(96)**
Atom Serial Number 54: SG **(97)**
Atom Serial Number 30: SG **(98)**
Atom Serial Number 208: SG **(99)**
Atom Serial Number 189: SG **(100)**

The distance between the center of the zinc atom **(91)** and:

the sulfur SG **(92)** is 2.409 A,
the sulfur SG **(93)** is 2.335 A,
the sulfur SG **(94)** is 2.346 A, and
the sulfur SG **(95)** is 2.285 A.

The distance between the center of the zinc atom **(96)** and:

the sulfur SG **(97)** is 2.450 A,
the sulfur SG **(98)** is 2.281 A,
the sulfur SG **(99)** is 2.298 A, and
the sulfur SG **(100)** is 2.305 A.

As can be seen by an examination of FIG. 15, two sets (set 1: (**(92)**,**(93)**,**(94)**,**(95)**) and set 2: (**(97)**,**(98)**,**(99)**,**(100)**)) of four SG atoms occupy the corners of two tetrahedrons

that each coordinate a zinc atom in the protein. Instead of having three coordinating histidine NE2 atoms **(40)**, **(42)**, and **(44)** as in 1AIY.PDB as shown in FIG. 9, 3M7K.PDB has four SG atoms from four cysteines that bind in a tetrahedral orientation relative to each bound zinc atom. There are two such zincs **(91)** and **(96)** tetrahedrally bound by cysteine SG atoms in 3M7K.PDB. The NDDPP **(1)** analysis of shared spatial volumes identifies several zinc binding sites, including binding sites of: three NE2 atoms shown in FIG. 9, two NE2 atoms and two SG atoms shown in FIG. 14, and four SG atoms as shown in FIG. 15.

Because processing by NDDPP **(1)** relies on occupancy and not orientation, atom configuration searches can be done without having to pre-orient data sets before comparison. Data sets can be reoriented relative to one another by matching any three volumes in space that satisfy matching occupancy criteria. In this manner, as little as three volumes can be mapped from one data set to another; coordinate and sizing transformations can be made to adjust the fit of the mapping.

The operations of NDDPP **(1)** as instructed by DEM **(8)** allow complex queries of data. The first embodiment describes how NDDPP **(1)** processes n-dimensional data sets to a data form that represents characterized spatial occupancy of n-dimensional data elements and then evaluates the characterized spatial occupancy of n-dimensional data elements to produce common spatial occupancy of n-dimensional data elements. Common spatial occupancy is then used to further characterize and evaluate additional data sets and to select data from data input to NDDPP **(1)**. For simplicity, the common spatial occupancy described in examples was three dimensional but one experienced in the field will understand that the common spatial occupancy of any number of dimensions can be evaluated.

In the examples, by the end of the execution of DEMS **(16)** SHIFT instructions, the data to be analyzed in the input data set has been converted to positive numbers. By changing the SHIFT value, the effect of volume grid placement can be assessed and adjusted. By changing the MULTFACTOR value, the volume grid is expanded or contracted and the effect of volume grid size can be assessed and adjusted. If elements in a data set or between data sets require changes in SHIFT or MULTFACTOR values in order to be identified, different SHIFT or MULTFACTOR values, whether specified or varied in an automated fashion, can be applied to input n-dimensional data using DEMI **(15)**.

By the end of the execution DEMS **(16)** MULTIPLY instructions, the columns to be analyzed in the input data set have been multiplied by a number that sizes the input data set. In the examples in the first embodiment, the MULTFACTOR for C7, C8, and C9 was variously 0.2 and 0.1. The sized data entries are then truncated. The effect of the sizing and truncating steps is to effectively select the size of a box in which to place the data elements; these data elements are atoms in the examples of the first embodiment. In all but one of the examples of the first embodiment, the size of the box was five angstroms on a side but the effect of sizing a box to be 10 angstroms on a side was shown in one example. The effect of sizing on the identification of common spatial occupancy can readily be explored by changing the MULTFACTOR value in DEMS **(16)**. By creating a new set of sized positive numbers in C7, C8, and C9 for every entry, a combined vector component can be made that represents the occupancy of a box in that vector space. With these transformations, complex data can be evaluated efficiently. If a set of selected elements occupy a small volume in one dataset but occupy a larger volume in another data set, as is the case in the example of 3DIM.PDB processing, the common spatial occupancy of n-dimensional data elements can still be identified by NDDPP **(1)** processing.

One skilled in the art will recognize that incremental changes in the values in the SHIFT and MULTIPLY instructions of DEMS **(16)** can be used to evaluate and optimize the output from DIPA **(2)**. It is understood that one skilled in the art can easily examine how any one set of DEMS **(16)** SHIFT and/or MULTIPLY instructions impact whether any one element is included in a volume by altering the SHIFT and/or MULTIPLY values incrementally to determine whether and how common spatial occupancy of any specified volume changes.

Second Embodiment

In the first embodiment of the present invention, a description of the processing by NDDPP **(1)** of input data sets with an intrinsic geometric component is provided. In the second embodiment, a description of the processing by NDDPP **(1)** of input data sets that have no intrinsic geometric component is provided.

As illustrated in FIG. 1 and FIG. 2 and described in FIG. 5 Step **(27)**, DS1 **(6)** and

DS2 **(7)** are input by USER **(5)** to DIPA **(2)** along with DEM **(8)**, and DSC **(14)**, if any. Since this is the first pass through the NDDPP **(1)**, there is no DSC **(14)** to input initially as described in FIG. 5 Step **(27)**. For the purposes of presenting an example of the Second Embodiment of the present invention, the patent title of the present invention, hereinafter referred to as DS1-W, will be used as an example of DS1 **(6)**; the patent title of the Matzusaki et al '683 patent, hereinafter referred to as DS2-W, will be used as an example of DS2 **(7)**; the instruction set DEMI-W will be used as a specific example of DEM **(8)** instruction set DEMI **(15)**; and the instruction set DEMS-W will be used as a specific example of DEM **(8)** instruction set DEMS **(16)**.

```
DS1-W:

N-DIMENSIONAL DATA PIPELINE PROCESSOR HAVING ADVANCED
FEATURES

DS2-W:

SIMILAR IMAGE RETRIEVING APPARATUS, THREE-DIMENSIONAL IMAGE
DATABASE APPARATUS AND METHOD FOR CONSTRUCTING THREE-
DIMENSIONAL IMAGE DATABASE
```

DS1-W and DS2-W data is not structured with the exception of punctuation such as commas, hyphens, line returns, spaces, and tabs.

As illustrated in FIG. 2 and described in FIG. 5 Step **(28)**, DEMI **(15)** of DEM **(8)** instructs DIPA-PA **(17)** to parse and attribute the input data sets. The example DEMI-W contains the following instructions for parsing DS1-W and DS2-W:

```
DEMI-W:

DATA PARSING FOR DS1-W and DS2-W:

DATA SEPARATION:
```

```
COMMAS,HYPHENS,LINE RETURNS,SPACES,TABS
C1:
  INPUT ENTRY NUMBER
C2:
  DATA ENTRY
```

DEMI-W instructs DIPA-PA **(17)**, for each entry in DS1-W AND DS2-W, to: read the characters in DS1-W; substitute a data delimiter for all punctuation such as commas, hyphens, line returns, spaces, and tabs; store the entry number in C1; and store the data entry in C2. For the purposes of this example, the data entry will be referred to as a "word."

After parsing according the DEMI-W, DS1-W has been parsed by DIPA-PA **(17)**:

```
1 N
2 DIMENSIONAL
3 DATA
4 PIPELINE
5 PROCESSOR
6 HAVING
7 ADVANCED
8 FEATURES
```

After parsing according the DEMI-W, DS2-W has been parsed by DIPA-PA **(17)**:

```
1 SIMILAR
2 IMAGE
3 RETRIEVING
4 APPARATUS
5 THREE
6 DIMENSIONAL
7 IMAGE
8 DATABASE
9 APPARATUS
10 AND
11 METHOD
```

12 FOR
13 CONSTRUCTING
14 THREE
15 DIMENSIONAL
16 IMAGE
17 DATABASE

DEMI-W next instructs, as described in FIG.5 Step **(28)**, the DIPA-PA **(17)** to attribute parsed DS1-W and DS2-W data. In this example, DEMI-W instructs DIPA-PA **(17)** to attribute parsed DS1-W and DS2-W data as follows:

```
DEMI-W (continued)

DATA ATTRIBUTES: FOR DS1-W and DS2-W:

IF C2 ENTRY EQUALS “ADVANCED”:      PUT  “6.14” IN C2 AND
                                    PUT  “3.26” IN C3
IF C2 ENTRY EQUALS “AND”:           PUT “11.02” IN C2 AND
                                    PUT  “5.10” IN C3
IF C2 ENTRY EQUALS “APPARATUS”:     PUT  “4.21” IN C2 AND
                                    PUT  “1.89” IN C3
IF C2 ENTRY EQUALS “CHAIRS”:        PUT “20.70” IN C2 AND
                                    PUT  “1.88” IN C3
IF C2 ENTRY EQUALS “CONSTRUCTING”: PUT “14.07” IN C2 AND
                                    PUT  “2.32” IN C3
IF C2 ENTRY EQUALS “DATA”:          PUT  “3.14” IN C2 AND
                                    PUT  “1.27” IN C3
IF C2 ENTRY EQUALS “DATABASE”:      PUT  “3.29” IN C2 AND
                                    PUT  “1.29” IN C3
IF C2 ENTRY EQUALS “DEVICE”:        PUT  “4.00” IN C2 AND
                                    PUT  “1.90” IN C3
IF C2 ENTRY EQUALS “DIMENSIONAL”:   PUT  “2.92” IN C2 AND
                                    PUT  “3.78” IN C3
IF C2 ENTRY EQUALS “FEATURES”:      PUT  “7.05” IN C2 AND
```

```
                                          PUT  “1.63” IN C3
    IF C2 ENTRY EQUALS “FOR”:             PUT “13.07” IN C2 AND
                                          PUT  “6.70” IN C3
    IF C2 ENTRY EQUALS “HAVING”:          PUT  “5.51” IN C2 AND
                                          PUT  “2.13” IN C3
    IF C2 ENTRY EQUALS “IMAGE”:           PUT  “9.84” IN C2 AND
                                          PUT  “1.62” IN C3
    IF C2 ENTRY EQUALS “INFORMATION”:     PUT  “3.00” IN C2 AND
                                          PUT  “1.20” IN C3
    IF C2 ENTRY EQUALS “METHOD”:          PUT “12.43” IN C2 AND
                                          PUT  “1.36” IN C3
    IF C2 ENTRY EQUALS “N”:               PUT  “1.97” IN C2 AND
                                          PUT  “7.01” IN C3
    IF C2 ENTRY EQUALS “NUMBER”:          PUT  “1.00 IN C2 AND
                                          PUT  “7.00” IN C3
    IF C2 ENTRY EQUALS “PIPELINE”:        PUT  “4.20” IN C2 AND
                                          PUT  “1.95” IN C3
    IF C2 ENTRY EQUALS “PROCESSOR”:       PUT  “4.10” IN C2 AND
                                          PUT  “1.90” IN C3
    IF C2 ENTRY EQUALS “RETRIEVING”:      PUT “10.72” IN C2 AND
                                          PUT  “2.45” IN C3
    IF C2 ENTRY EQUALS “SIMILAR”:         PUT  “8.56” IN C2AND
                                          PUT  “3.67” IN C3
    IF C2 ENTRY EQUALS “THREE”:           PUT  “1.38” IN C2 AND
                                          PUT  “7.30” IN C3
```

One skilled in the art will recognize that any database of words, word meaning, and other attributes of character information can be attributed to the parsed entries as instructed by DEMI **(15)**. One skilled in the art will recognize that DEMI **(15)** can extract and use any information from or feature of input data, such as the most common character in data, to parse data.

In this example of the second embodiment of the present invention, the DEMI-W instructs the DIPA-PA **(17)** to both replace an entry and add an entry. DEMI-W instructs the

DIPA-PA **(17)** to make PADS1 **(18)** and PADS2 **(19)**. In this example, the example of PADS1 **(18)** is hereinafter referred to as PADS1-W and the example of PADS1 **(19)** is hereinafter referred to as PADS2-W. Parsed and attributed PADS1-W and PADS2-W are output to DIPA-SST **(20)**.

PADS1-W:

1	1.97	7.01
2	2.92	3.78
3	3.14	1.27
4	4.20	1.95
5	4.10	1.90
6	5.51	2.13
7	6.14	3.26
8	7.05	1.63

PADS2-W:

1	8.56	3.67
2	9.84	1.62
3	10.72	2.45
4	4.21	1.89
5	1.38	7.30
6	2.92	3.78
7	9.84	1.62
8	3.29	1.29
9	4.21	1.89
10	11.02	5.10
11	12.43	1.36
12	13.07	6.70
13	14.07	2.32
14	1.38	7.30
15	2.92	3.78
16	9.84	1.62
17	3.29	1.29

In this example, the word in C2 is substituted with a number whose integer portion is the same for all synonyms of the substituted word, and a number is added in C3 whose integer portion is the same for all words in the same word class as the substituted word.

In this example, “DATA” is mapped to 3.14 in C2, and “DATABASE” is mapped to 3.29 in C2. “DATA” and “DATABASE” have a similar meaning so these words are replaced by numbers that are the same to the left of the decimal point but are different to the right of the decimal point. Similarly in C2: “PROCESSOR” is mapped to 4.10, “PIPELINE” is mapped to 4.20, and “APPARATUS” is mapped to 4.21.

In this example, word entries that are nouns, verbs, adjectives, adverbs and conjunctions are assigned different numbers in C3, nouns being assigned a “1” plus a fraction, verbs assigned a “2” plus a fraction, adjectives assigned a “3” plus a fraction, adverbs assigned a “4” plus a fraction, conjunctions assigned a “5” plus a fraction, prepositions a “6” plus a fraction, and numbers assigned a “7” plus a fraction. For the purposes of this example, words which are used as adjectives that are also nouns have also been assigned as nouns, such as “data”, “image”, and “database”.

In this example of the second embodiment, DS1-W and DS2-W entering DIPA-PA **(17)** are parsed, in this example delimited by punctuation; attributed, in this example replaced with numbers representing the word and an attribute of the word; and output as PADS1-W and PADS2-W to DIPA-SST **(20)**. The parsed and attributed information, PADS1-W and PADS2-W are passed to DIPA-SST **(20)** as illustrated in FIG. 2 and described in FIG. 5 Step **(29)**. The DIPA-SST **(20)** shifts, sizes, and truncates PADS1-W and PADS2-W as instructed by DEMS **(16)**. An example DEMS **(16)** is DEMS-W.

```
DEMS-W:

FOR ALL INPUT PADS1-W and PADS2-W ENTRIES:

C1 TO C1

MULTFACTOR = 1.0
```

```
SHIFT COMMANDS:
  NO SHIFT (ADD ZERO TO ALL ENTRIES)

SIZE COMMANDS:
  NO SIZE (MULTFACTOR = 1.0)

TRUNCATE COMMANDS:
  TRUNCATE THE CURRENT VALUE OF C2 AFTER THE DECIMAL
    PLACE AND PUT THE INTEGER RESULT IN C2.
  TRUNCATE THE CURRENT VALUE OF C3 AFTER THE DECIMAL
    PLACE AND PUT THE INTEGER RESULT IN C3.
```

DEMI-W instructs the DIPA-SST **(20)** to make PDS1-W, an example of PDS1 **(9)** and PDS2-W, an example of PDS2 **(10)**. Parsed, attributed, shifted, sized, and truncated PDS1-W and PDS2-W are output to MOPA **(3)** as illustrated in FIG. 1 and FIG. 2 and described in FIG. 5 Step **(30)**.

```
PDS1-W:
  1        1        7
  2        2        3
  3        3        1
  4        4        1
  5        4        1
  6        5        2
  7        6        3
  8        7        1

PDS2-W:
  1        8        3
  2        9        1
  3       10        2
  4        4        1
  5        1        7
  6        2        3
```

```
7        9        1
8        3        1
9        4        1
10       11       5
11       12       1
12       13       6
13       14       2
14       1        7
15       2        3
16       9        1
17       3        1
```

As illustrated in FIG. 1 and FIG. 3 and described in FIG. 6 Step **(31)**, PDS1 **(9)** and PDS2 **(10)** are input to MOPA **(3)** along with DEM **(8)** instruction set DEMM **(21)** and MC **(13)**, if any. Since this is the first pass through the NDDPP **(1)**, there is no MC **(13)** to input initially as described in FIG. 5 Step **(31)**. FIG. 6, Step **(32)** describes the use of the integer values of PDS1 **(9)** and PDS2 **(10)** to make the matrix indices M1 **(11)** and M2 **(12)**. In this example, M1-W will be used as a specific example of M1 **(11)** and M2-W will be used as a specific example of M2 **(12)**. As described in FIG. 6 Step **(32)**, the integer values in columns of the PDS1-W are used make M1-W and the integer values in the columns of PDS2-W are used to make M1-W. In this example, for every entry in PDS1-W, the value of C1 becomes the value of the first index in M1-W; the value of C2 becomes the value of the second index in M1-W; and the value of C3 becomes the value of the third index in M1-W. Similarly, for every entry in PDS2-W, the value of C1 becomes the value of the first index in M2-W; the value of C2 becomes the value of the second index in M2-W; and the value of C3 becomes the value of the third index in M2-W.

```
M1-W:
M1-W(  1,  1,  7 )
M1-W(  2,  2,  3 )
M1-W(  3,  3,  1 )
M1-W(  4,  4   1 )
M1-W(  5,  4,  1 )
```

```
M1-W(  6,  5,  2 )
M1-W(  7,  6,  3 )
M1-W(  8,  7,  1 )

M2-W:
M2-W(  1,  8,  3 )
M2-W(  2,  9,  1 )
M2-W(  3, 10,  2 )
M2-W(  4,  4,  1 )
M2-W(  5,  1,  7 )
M2-W(  6,  2,  3 )
M2-W(  7,  9,  1 )
M2-W(  8,  3,  1 )
M2-W(  9,  4,  1 )
M2-W( 10, 11,  5 )
M2-W( 11, 12,  1 )
M2-W( 12, 13,  6 )
M2-W( 13, 14,  2 )
M2-W( 14,  1,  7 )
M2-W( 15,  2,  3 )
M2-W( 16,  9,  1 )
M2-W( 17,  3,  1 )
```

FIG. 16 is a view showing a visualization of MOPA **(3)** output of M1-W and M2-W. In FIG. 16, hexagons represent M1-W indices and the circles represent M2-W indices where the values of the first, second and third indices are used as the x, y, z coordinates:

M1-W (1,1,7) is represented by a hexagon **(101)** at coordinate x=1, y=1, and z=7

M1-W (2,2,3) is represented by a hexagon **(102)** at coordinate x=2, y=2, and z=3

M1-W (3,3,1) is represented by a hexagon **(103)** at coordinate x=3, y=3, and z=1

M1-W (4,4,1) is represented by a hexagon **(104)** at coordinate x=4, y=4, and z=1

M1-W (5,4,1) is represented by a hexagon **(105)** at coordinate x=5, y=4, and z=1

M1-W (6,5,2) is represented by a hexagon **(106)** at coordinate x=6, y=5, and z=2

M1-W (7,6,3) is represented by a hexagon **(107)** at coordinate x=7, y=6, and z=3

M1-W (8,7,1) is represented by a hexagon **(108)** at coordinate x=8, y=7, and z=1

M2-W (1,8,3) is represented by a circle **(109)** at coordinate x=1, y=8, and z=3

M2-W (2,9,1) is represented by a circle **(110)** at coordinate x=2, y=9, and z=1

M2-W (3,10,2) is represented by a circle **(111)** at coordinate x=3, y=10, and z=2

M2-W (4,4,1) is represented by a circle **(112)** at coordinate x=4, y=4, and z=1

M2-W (5,1,7) is represented by a circle **(113)** at coordinate x=5, y=1, and z=7

M2-W (6,2,3) is represented by a circle **(114)** at coordinate x=6, y=2, and z=3

M2-W (7,9,1) is represented by a circle **(115)** at coordinate x=7, y=9, and z=1

M2-W (8,3,1) is represented by a circle **(116)** at coordinate x=8, y=3, and z=1

M2-W (9,4,1) is represented by a circle **(117)** at coordinate x=9, y=4, and z=1

M2-W (10,11,5) is represented by a circle **(118)** at coordinate x=10, y=11, and z=5

M2-W (11,12,1) is represented by a circle **(119)** at coordinate x=11, y=12, and z=1

M2-W (12,13,6) is represented by a circle **(120)** at coordinate x=12, y=13, and z=6

M2-W (13,14,2) is represented by a circle **(121)** at coordinate x=13, y=14, and z=2

M2-W (14,1,7) is represented by a circle **(122)** at coordinate x=14, y=1, and z=7

M2-W (15,2,3) is represented by a circle **(123)** at coordinate x=15, y=2, and z=3

M2-W (16,9,1) is represented by a circle **(124)** at coordinate x=16, y=9, and z=1

M2-W (17,3,1) is represented by a circle **(125)** at coordinate x=17, y=3, and z=1

Among other observations, it can be seen from an examination of FIG. 16 and the discussion of the present example of the second embodiment, that:

hexagon **(101)**, circle **(113),** and circle **(122)** are aligned;

hexagon **(102)**, circle **(114)**, and circle **(123)** are aligned;

hexagon **(103)**, circle **(116)**, and circle **(125)** are aligned;

hexagon **(101)** is derived from the processing of the word “N”;

circles **(113)** and **(122)** are derived from the processing of the word “three”;

hexagon **(102)**, circle **(114)**, and circle **(123)** are derived from the processing of the word “dimensional”;

hexagon **(103)** is derived from the processing of the word “DATA”; and

circles **(116)** and **(125)** are derived from the processing of the word “DATABASE”.

M1 **(11)** and M2 **(12)** output contain information on the characterized spatial

occupancy of n-dimensional data elements in DS1 **(6)** and DS2 **(7)** input as illustrated by the example of characterized word data in FIG. 16.

M1-W and M2-W are output by MOPA-MM **(22)** to USER **(5)** and also input to MOPA-CM **(23)** along with DEMM **(21)** as illustrated in FIG. 3 and described in FIG. 6 Steps **(32)** and **(33)**. There is no MC **(13)** input to the MOPA-CM **(23)** at this point in the example of the second embodiment since MC **(13)** is an output of NDDPP **(1)** and it is the first pass of data through NDDPP **(1)**. MOPA-CM **(23)** makes MC **(13)** as instructed by DEMM **(21)** and described in FIG. 6 Step **(33)**. DEMM **(21)** can instruct MOPA-CM **(23)** to output any combination of information related to shared indices of M1 **(11)** and M2 **(12)** and whether and how to use previously produced MC **(13)** in new MC **(13)** production. In this example of the second embodiment of the present invention, DEMM-W will be used as a specific example of DEMM **(21)**.

```
DEMM-W:

FOR ALL M1-W AND M2-W MATRIX INDICES:

IF:
  M1-W SECOND INDEX VALUE EQUALS THE
    M2-W SECOND INDEX VALUE
THEN MAKE MC-W MATRIX INDICES:
  PUT M1-W  FIRST INDEX VALUE IN MC-W  FIRST INDEX
  PUT M2-W  FIRST INDEX VALUE IN MC-W SECOND INDEX AND
  PUT M1-W SECOND INDEX VALUE IN MC-W  THIRD INDEX.
```

For example, certain M1-W and M2-W indices share the same number, “1”, as the value of the second index:

```
M1-W(  1,  1,  7 )
M2-W(  5,  1,  7 )
M2-W( 14,  1,  7 )
```

The following MC-W indices are made representing the shared number 1 as the value of the second index of M1-W and M2-W:

```
MC-W(  1,  5,  1 )
MC-W(  1, 14,  1 )
```

M1-W and M2-W share the same number, “2” , as the value of the second index.

```
M1-W(  2,  2,  3 )
M2-W(  6,  2,  3 )
M2-W( 15,  2,  3 )
```

The following MC-W indices are made representing the shared word “2” as the value of the second index of M1-W and M2-W:

```
MC-W(  2,  6,  2 )
MC-W(  2, 15,  2 )
```

M1-W and M2-W share the same number,“3”, as the value of the second index.

```
M1-W(  3,  3,  1 )
M2-W(  8,  3,  1 )
M2-W( 17,  3,  1 )
```

The following MC-W indices are made representing the shared word “3” as the value of the second index of M1-W and M2-W:

```
MC-W(  3,  8,  3 )
MC-W(  3, 17,  3 )
```

M1-W and M2-W share the same number, “4” , as the value of the second index:

```
M1-W(  4,  4,  1 )
M1-W(  5,  4,  1 )
M2-W(  4,  4,  1 )
M2-W(  9,  4,  1 )
```

The following MC-W indices are made representing the shared word “3” as the value of the second index of M1-W and M2-W:

```
MC-W(  4,  4,  4 )
MC-W(  4,  9,  4 )
MC-W(  5,  4,  4 )
MC-W(  5,  9,  4 )
```

It can be seen from an inspection of the M1-W and M2-W indices that values of 5-7 are not found as the value of the second index of M2-W and values of 8-14 are not found as the value of the second index of M1-W:

```
M1-W(  6,  5,  2 )
M1-W(  7,  6,  3 )
M1-W(  8,  7,  1 )
M2-W(  1,  8,  3 )
M2-W(  2,  9,  1 )
M2-W(  7,  9,  1 )
M2-W( 16,  9,  1 )
M2-W(  3, 10,  2 )
M2-W( 10, 11,  5 )
M2-W( 11, 12,  1 )
M2-W( 12, 13,  6 )
M2-W( 13, 14,  2 )
```

However, there are instances where M1-W and M2-W share values in the third indices but not in the second indices:

```
M1-W(  8,  7,  1 )
M2-W(  2,  9,  1 )
M2-W(  7,  9,  1 )
M2-W( 11, 12,  1 )
M2-W( 16,  9,  1 )
M1-W(  6,  5,  2 )
M2-W(  3, 10,  2 )
M2-W( 13, 14,  2 )
M1-W(  7,  6,  3 )
M2-W(  1,  8,  3 )
```

When matrix elements that do not share common second index values are ordered by third index value, it can be seen that although the words with meanings in categories 5-14 did not occur in both M1-W and M2-W, as indicated by no shared values in the second indices of M1-W and M2-W, there were shared values in index 3. Although they did not share a similar meaning, nouns, indicated by a "1" as the value in the third matrix index position; verbs, indicated by a "2" as the value in the third matrix index position; and adjectives, as indicated by a "3" as the value in the third matrix position, were present in both M1-W and M2-W.

In this example of the first embodiment of the present invention, processed M1-W, M2-W, and MC-W are output by MOPA **(3)** as illustrated in FIG. 3 and described in FIG. 6 Step **(34)**. M1-W, M2-W, and MC-W are output to USER **(5)**. MC-W is also output to FOPA **(4)** and MOPA-CM **(23)**:

```
MC-W(  1,  5,  1 )
MC-W(  1, 14,  1 )
MC-W(  2,  6,  2 )
MC-W(  2, 15,  2 )
MC-W(  3,  8,  3 )
MC-W(  3, 17,  3 )
MC-W(  4,  4,  4 )
MC-W(  4,  9,  4 )
```

```
MC-W(  5,  4,  4 )
MC-W(  5,  9,  4 )
```

FIG. 17 is a view showing a visualization of MOPA **(3)** output of MC-W. FIG. 17 shows the ten points whose x, y, and z coordinates are the values of the first, second, and third MC-W indices. FIG. 17 shows:

MC-W (1,5,1) is represented by a circle **(126)** at coordinate x=1, y=5 and z=1 indicating common spatial occupancy of a number assignment (index 3 in MC-W), "N" (entry 1 in DS1-W represented by index 1 in MC-W), and "THREE" (entry 5 in DS2-W represented by index 2 in MC-W);

M2-W (1,14,1) is represented by a circle **(127)** at coordinate x=1, y=14 and z=1 indicating common spatial occupancy of a number assignment (index 3 in MC-W), "N" (entry 1 in DS1-W represented by index 1 in MC-W), and "THREE" (entry 14 in DS2-W represented by index 2 in MC-W);

M2-W (2, 6, 2) is represented by a circle **(128)** at coordinate x=2, y=6 and z=2 indicating common spatial occupancy of an adjective assignment (index 3 in MC-W), "DIMENSIONAL" (entry 2 in DS1-W represented by index 1 in MC-W), and "DIMENSIONAL" (entry 6 in DS2-W represented by index 2 in MC-W);

M2-W (2, 15, 2) is represented by a circle **(129)** at coordinate x=2, y=15 and z=2 indicating common spatial occupancy of an adjective assignment (index 3 in MC-W), "DIMENSIONAL" (entry 2 in DS1-W represented by index 1 in MC-W), and "DIMENSIONAL" (entry 15 in DS2-W represented by index 2 in MC-W);

M2-W (3, 8, 3) is represented by a circle **(130)** at coordinate x=3, y=8 and z=3 indicating common spatial occupancy of a noun assignment (index 3 in MC-W), "DATA" (entry 3 in DS1-W represented by index 1 in MC-W), and

"DATABASE" (entry 8 in DS2-W represented by index 2 in MC-W);

M2-W (3, 17, 3) is represented by a circle **(131)** at coordinate x=3, y=17 and z=3 indicating common spatial occupancy of a noun assignment (index 3 in MC-W), "DATA" (entry 3 in DS1-W represented by index 1 in MC-W), and "DATABASE" (entry 17 in DS2-W represented by index 2 in MC-W);

M2-W (4, 4, 4) is represented by a circle **(132)** at coordinate x=4, y=4 and z=4 indicating common spatial occupancy of a noun assignment (index 3 in MC-W), "PIPELINE" (entry 4 in DS1-W represented by index 1 in MC-W), and "APPARATUS" (entry 4 in DS2-W represented by index 2 in MC-W);

M2-W (4, 9, 4) is represented by a circle **(134)** at coordinate x=4, y=9 and z=4 indicating common spatial occupancy of a noun assignment (index 3 in MC-W), "PIPELINE" (entry 4 in DS1-W represented by index 1 in MC-W), and "APPARATUS" (entry 9 in DS2-W represented by index 2 in MC-W); and

M2-W (5, 4, 4) is represented by a circle **(133)** at coordinate x=5, y=4 and z=4 indicating common spatial occupancy of a noun assignment (index 3 in MC-W), "PROCESSOR" (entry 5 in DS1-W represented by index 1 in MC-W), and "APPARATUS" (entry 4 in DS2-W represented by index 2 in MC-W);

M2-W (5, 9, 4) is represented by a circle **(135)** at coordinate x=5, y=9 and z=4 indicating common spatial occupancy of a noun assignment (index 3 in MC-W), "PROCESSOR" (entry 5 in DS1-W represented by index 1 in MC-W), and "APPARATUS" (entry 9 in DS2-W represented by index 2 in MC-W).

MOPA-CM **(23)** takes the characterized spatial occupancy of n-dimensional data elements represented by M1 **(11)** and M2 **(12)**, an example of which is shown in FIG. 16, and outputs common spatial occupancy of n-dimensional data elements represented by MC

(13), an example of which is shown in FIG. 17. FIG. 17 represents a compact visual representation of common spatial occupancy of characterized words in input data sets.

FIG. 7 is a flowchart showing each step of the processing procedure of FOPA **(4)** and outlines how the example data is processed through FOPA **(4)** of the NDDPP **(1)** as illustrated in FIG. 1 and FIG. 4. FIG. 4 illustrates the input of DS1 **(6)** and DS2 **(7)** to FOPA-AC **(26)** and MC **(13)** to FOPA-CC **(24)**. The example input data sets, DS1-W and DS2-W are input to FOPA-AC **(26)** and MC-W **(13)** is input into FOPA-CC **(24)** as illustrated in FIG. 4 and described in FIG. 7 Step **(35).**

FOPA-CC **(24)** outputs selection criteria SCF **(25)** based on MC **(13)** as illustrated in FIG. 4 and described in FIG. 7 Step **(36)**. In this example of the second embodiment, an example of SCF **(25)** is referred to as SCF-W and is output as:

```
SCF-W:

FOR ALL MC-W MATRIX INDICES:

FOR ALL PARSED DS1-W INPUT LINES THAT HAVE AN ENTRY NUMBER
  IN C1 THAT MATCHES THE VALUE OF INDEX 1 IN MC-W INDICES:
    PUT THE WORD IN C2 INTO DSC-W
```

As illustrated in FIG. 4 and described in FIG. 7 Step **(37)** and using SCF-W, FOPA-AC **(26)** outputs DSC-W, an example of DSC **(14)**:

```
N
DIMENSIONAL
DATA
PIPELINE
PROCESSOR
```

DSC-W contains the words “N DIMENSIONAL DATA PIPELINE PROCESSOR” which, in this example, happen to be words that are contiguous in DS1-W. For the purposes of this example, DSC-W is output to DIPA-PA **(17)** and to USER **(5)** as illustrated in FIG. 1

and FIG. 4 and described in FIG. 7 Step **(38)**.

In order to illustrate an example of DSC **(14)** feedback to DIPA **(2)**, DSC-W output from the analysis of DS1-W and DS2-W will be used as an example of DSC **(14)** that is re-input to NDDPP **(1)** and referred to as DS1-W2. A new input data set, DS2-W2, will be used as an example of DS2 **(7)**:

```
DS2-W2:

METHOD FOR CONSTRUCTING THREE DIMENSIONAL CHAIRS
```

After parsing according the DEMI-W, DS1-W2 has been parsed by DIPA-PA **(17)**:

```
1 N
2 DIMENSIONAL
3 DATA
4 PIPELINE
5 PROCESSOR
```

After parsing according the DEMI-W, DS2-W2 has been parsed by DIPA-PA **(17)**:

```
1 METHOD
2 FOR
3 CONSTRUCTING
4 THREE
5 DIMENSIONAL
6 CHAIRS
```

DEMI-W is used to make PADS1-W2 in DIPA-PA **(17)**:

```
1       1.97     7.01
2       2.92     3.78
3       3.14     1.27
4       4.20     1.95
```

```
5        4.10      1.90
```

DEMI-W is used to make PADS2-W2 in DIPA-PA **(17)**:

```
1        12.43     1.36
2        13.07     6.70
3        14.07     2.32
4         1.38     7.30
5         2.92     3.78
6        20.70     1.88
```

DEMS-W is used to make PDS1-W2 in DIPA-SST **(20)**:

```
1         1         7
2         2         3
3         3         1
4         4         1
5         4         1
```

DEMS-W is used to make PDS2-W2 in DIPA-SST **(20)**:

```
1        12         1
2        13         6
3        14         2
4         1         7
5         2         3
6        20         1
```

M1-W2 is made from PDS1-W2 in MOPA-MM **(22)**:

```
M1-W2 (  1,  1,  7)
M1-W2 (  2,  2,  3)
M1-W2 (  3,  3,  1)
```

```
M1-W2 (  4,  4   1)
M1-W2 (  5,  4,  1)
```

M2-W2 is made from PDS2-W2 in MOPA-MM **(22)**:

```
M2-W2 (  1, 12,  1 )
M2-W2 (  2, 13,  6 )
M2-W2 (  3, 14,  2 )
M2-W2 (  4,  1,  7 )
M2-W2 (  5,  2,  3 )
M2-W2 (  6, 20,  1 )
```

MC-W2 is made using DEMM-W in MOPA-CM **(23)**:

```
MC-W2(  1,  4,  1 )
MC-W2(  2,  5,  2 )
```

MC-W2 is output to MOPA-CM **(23)**, USER **(5)**, and FOPA **(4)**. If MC-W2 is fed back to MOPA-CM **(23)**, the MC-W2 indices could be used to eliminate MC **(13)** entries being generated by MOPA-CM **(23)** that did not contain similar matrix indices derived from processing of input data containing words with the same attribute values. In this way, volumes with mapped points representing this phrase can rapidly be identified graphically and computationally.

Using SCF-W, a second example of DSC **(14)**, DSC-W2 is made:

```
N
DIMENSIONAL
```

The words “N” and “DIMENSIONAL” in DSC-W2 represent words in input DS1-W, DS2-W, and DS2-W2.

MC **(13)** output can be visualized to see the common spatial occupancy of the words,

as shown in FIG. 17. Once symbolic data is converted to n-dimensional data representing common spatial occupancy, the n-dimensional data can itself be processed similar to data in the first embodiment in order to look at common spatial occupancy of volumes and higher dimensional spaces. One skilled in the art would conclude that NDDPP **(1)** as described in the second embodiment can analyze any number of data sets with any number of dimensions. Data input to, and processed by, NDDPP **(1)** need not be input as files, but can be continuously streamed through the NDDPP **(1)**. By extension, matrix indices can be formed for any number of input data sets that represent the sharing of common elements in the input data, here illustrated by example as words. One skilled in the art will conclude that an advantage of the present invention is that no predetermined search information needs to be supplied in order to find patterns in data and across data sets.

As an alternative to assigning specific values to specific words as instructed by DEMI-W, unique characters or symbols can be assigned a unique identifier in their order of appearance. In this alternative method, the same character or number receives the same assignment so that new numbers are assigned only when a new character is seen on the input. This method of assignment by order of occurrence is particularly useful when the meaning of the characters or symbols are not known, and it is not known how characters or symbols in the data are grouped. Character assignment made by order of unique occurrence can then be used in DEMI **(15)** and used by DIPA-PA **(17)** to assign a unique number to unattributed input data. Thus, it is not necessary to delimit character or symbolic data in order for NDDPP **(1)** to operate on data as discussed in the second embodiment. Unstructured data or symbols can be parsed by any method that delimits a data element. Data elements or groups of data elements in unstructured data that are identified by DSC **(14)** can be fed back to DIPA **(2)**.

An advantage of the present invention is that once data elements are grouped, either by assignment or by feedback of DSC **(14)** to NDDPP **(1)**, numbers may be assigned in such a manner that the truncation instruction by DEMS **(16)** effectively places members of the group in a common spatial occupancy. The ability to use the NDDPP **(1)** to find common spatial occupancy of n-dimensional data elements which can be used to structure data is a unique feature of the present invention. It is not necessary to pre-select particular n-tuples or particular features of data in order to use the n-dimensional data pipeline processor to

indentify relationships between components of entered n-dimensional data.

Third Embodiment

In the example of the first embodiment of the present invention, NDDPP **(1)** processing of input data sets with a geometric component was described. In the example of the second embodiment of the present invention, NDDPP **(1)** processing of input data sets with no intrinsic geometric component was described. In the example of the third embodiment of the present invention, NDDPP **(1)** processing of image data will be described.

As illustrated in FIG. 1 and FIG. 2 and described in FIG. 5 Step **(27)**, DIPA **(2)** inputs DS1 **(6)**, DS2 **(7)**, and DEM **(8)** from USER **(5),** along with DSC **(14)** from FOPA **(4)**, if any. Since this is the first pass through the NDDPP **(1)**, there is no DSC **(14)** to input initially as described in FIG. 5 Step **(27)**. For the purposes of the example of the third embodiment of the present invention, data to be processed is in the form of a stream of pixels that form a 2D image that are input as rows of pixels, i.e., a scan line. Each pixel may have any number of primary or compound attributes, such as area, color, contrast texture, or density. To provide a simple example, the following data will be used as an example of DS1 **(6)**, hereinafter referred to as DS1-I:

```
 1.75  1.64  1.82  1.33  1.78  1.78  1.83  1.16  1.52  1.71
 1.20  1.19 20.09 20.14  1.56  1.48  1.56  1.56  3.23  1.46
 1.03 60.48  1.61  1.04  1.54  1.32  1.66  1.48  1.21  1.33
80.93  1.60  1.72  1.43  1.60  1.24  1.91  1.41  1.27  1.37
 1.24  1.58  1.70  1.06  1.60  1.63  1.15  1.58  1.81  1.74
```

The following data will be used as an example of DS2 **(7)**, hereinafter referred to as DS2-I:

```
 1.89   1.25  1.90  1.36  1.00  1.17 80.15  1.36  1.43  1.67
 1.88   1.20  1.46  1.65  1.21  1.11  1.33 60.05  1.30  2.54
 1.41   1.51  1.16  1.09  1.43  1.38  1.41  1.46 20.95  1.19
```

```
1.75   1.95  1.91  1.95  3.30  1.30  1.57  1.27 20.03  1.65
1.19   1.32  1.61  1.17  1.93  1.92  1.18  1.61  1.17  1.38
```

The numbers in DS1-I and DS2-I are representative of a property of a pixel. The numbers in a line represent a scan line of pixels. For the purposes of this example, a single pixel value is delimited by tabs and line returns. The values attributed to the pixel and delimited by tabs and lines can represent any one of many attributes which are commonly associated with pixel data.

As illustrated on FIG. 2, DEMI **(15)** is input by DIPA-PA **(17)**. DEMI **(15)** contains instructions for the parsing and attribution of the input data sets. For the purposes of this example, DEMI-I will be used as an example of the DEMI **(15)**. DEMI-I contains the following information for formatting DS1-I and DS2-I:

```
DEMI-I:

DATA PARSING:
  5 x 5 PIXEL SECTORS:

DATA SEPARATION:
  TABS AND LINE RETURNS
```

DEMI-I instructs DIPA-PA **(17)**, for each entry in DS1-I and DS2-I to:

parse the value associated with each pixel using the example data delimiter of tabs and line returns;

store the delimited number sequentially; and

reformat the parsed and stored data into sectors that represent elements of a spatial sector. For the purposes of the example, the spatial grid is a 5 x 5 pixel sector and the attribute assigned to the pixel will be a number that can represent any feature of the pixel, for example, color. In this example of the third embodiment, data sets given below only contain five scan lines and there are only two sectors each derived from DS1-I and DS2-I. If the scan line was longer or there were more than five lines in the input data, it is understood by one

skilled in the art that there would be additional sectors in DS1-I and DS2-I.

The DS1-I first 5x5 pixel sector, hereinafter referred to as DS1-IS1, consists of the first set of five values in each of the first five scan lines in DS1-I. The DS1-I second 5x5 pixel sector, hereinafter referred to as DS1-IS2, consists of the second set of five values in each of the first five scan lines in DS1-I. The DS2-I first 5x5 pixel sector, hereinafter referred to as DS2-IS1, consists of the first set of five values in each of the first set of five scan lines in DS2-I. The DS2-I second 5x5 pixel sector, hereinafter referred to as DS2-IS2, consists of the second set of five values in each of the first set of five scan lines in DS2-I.

```
DS1-IS1:
 1.75   1.64   1.82    1.33   1.78
 1.20   1.19  20.09   20.14   1.56
 1.03  60.48   1.61    1.04   1.54
80.93   1.60   1.72    1.43   1.60
 1.24   1.58   1.70    1.06   1.60

DS1-IS2:
 1.78   1.83   1.16    1.52   1.71
 1.48   1.56   1.56    3.23   1.46
 1.32   1.66   1.48    1.21   1.33
 1.24   1.91   1.41    1.27   1.37
 1.63   1.15   1.58    1.81   1.74

DS2-IS1:
 1.89   1.25   1.90    1.36   1.00
 1.88   1.20   1.46    1.65   1.21
 1.41   1.51   1.16    1.09   1.43
 1.75   1.95   1.91    1.95   3.30
 1.19   1.32   1.61    1.17   1.93

DS2-IS2:
 1.17  80.15   1.36    1.43   1.67
 1.11   1.33  60.05    1.30   2.54
```

```
1.38   1.41   1.46   20.95   1.19
1.30   1.57   1.27   20.03   1.65
1.92   1.18   1.61    1.17   1.38
```

DEMI-I next instructs, as described in FIG. 5 Step **(28)**, the DIPA-PA **(17)** to attribute parsed DS1-I and DS2-I data. In this example, DEMI-I instructs DIPA-PA **(17)** to not attribute further the parsed DS1-I and DS2-I data as follows:

```
DATA ATTRIBUTES:
  PIXEL VALUE
```

In this example, for simplicity, no additional attributes will be assigned to the pixel, but one skilled in the art will understand that any number of attributes can be assigned to the pixel and the present example provides the basis for expanding the dimensionality of the input data prior to further analysis.

DEMI-I instructs the DIPA-PA **(17)** to make PADS1 **(18)** and PADS2 **(19)** without further attributing the data. In this example, the example of PADS1 **(18)** is hereinafter referred to as PADS1-I and is comprised of DS1-IS1 and DS1-IS2; and the example of PADS2 **(19)** is hereinafter referred to as PADS2I and is comprised of DS2- S1 and DS2-IS2. Parsed and attributed PADS1-I and PADS2-I are output to DIPA-SST **(20)**.

In this example, the parsed and attributed information, PADS1-I and PADS2-I are passed to DIPA-SST **(20)** as illustrated in FIG. 2 and described in FIG. 5 Step **(29)**. The DIPA-SST **(20)** shifts, sizes, and truncates PADS1-I and PADS2-I as instructed by DEMS **(16)**. The example DEMS **(16)** in this example is hereinafter referred to as DEMS-I.

```
DEMS-I:

MULTFACTOR = 0.1

FOR ALL INPUT PADS1-I and PADS2-I ENTRIES:
```

```
SHIFT COMMANDS:
  NO SHIFT = ADD ZERO TO ALL ENTRIES

SIZE COMMANDS:
  MULTIPLY ALL ENTRIES BY MULTFACTOR

TRUNCATE COMMANDS:
  TRUNCATE THE CURRENT VALUE OF EVERY PIXEL AFTER THE
    DECIMAL PLACE AND
      PUT THE INTEGER RESULT BACK IN THAT PIXEL.
```

DEMI-I instructs the DIPA-SST **(20)** to make PDS1 **(9)** and PDS2 **(10)** by sizing and truncating the values of the pixels. In this example, the example of PDS1 **(9)** is hereinafter referred to as PDS1-I and the example of PDS2 **(10)** is hereinafter referred to as PDS2-I. Parsed, attributed, shifted, sized, and truncated PDS1-I and PDS2-I are output to MOPA **(3)** as illustrated in FIG. 1 and FIG. 2 and described in FIG. 5 Step **(30)**.

The PDS1-I first processed 5x5 pixel sector, PDS1-IS1, consists of the shifted, sized, and truncated number values from DS1-IS1. The PDS1-I second processed 5x5 pixel sector, PDS1-IS2, consists of the shifted, sized, and truncated number values from DS1-IS2. The PDS2-I first processed 5x5 pixel sector, PDS2-IS1, consists of the shifted, sized, and truncated number values from DS2-IS1. The PDS2-I second processed 5x5 pixel sector, PDS2-IS2, consists of the shifted, sized, and truncated number from DS2-IS2.

```
PDS1-IS1:
  0     0     0     0     0
  0     0     2     2     0
  0     6     0     0     0
  8     0     0     0     0
  0     0     0     0     0

PDS1-IS2:
  0     0     0     0     0
```

```
0    0    0    0    0
0    0    0    0    0
0    0    0    0    0
0    0    0    0    0
```

```
PDS2-IS1:
0    0    0    0    0
0    0    0    0    0
0    0    0    0    0
0    0    0    0    0
0    0    0    0    0
```

```
PDS2-IS2:
0    8    0    0    0
0    0    6    0    0
0    0    0    2    0
0    0    0    2    0
0    0    0    0    0
```

FIG. 18 shows a view of a representation of the processed 5x5 pixel sectors PDS1-IS1 **(136)**, PDS1-IS2 **(137)**, PDS2-IS1 **(138)** and PDS2-IS2 **(139)**. In FIG. 18, PDS1-IS1 **(136)** and PDS2-IS2 **(139)** each have four circles whose size is proportional to the value of the attribute associated with that pixel. It can be readily seen from an examination of FIG. 18 that the pattern of the circles in PDS1-IS1 **(136)** and PDS2-IS2 **(139)** is the same. PDS1-IS1 **(136)** and PDS2-IS2 **(139)** contain the same number of non-zero pixels and the positions of the non-zero pixels are related by a rotation and translation. Both PDS1-IS2 **(137)** and PDS2-IS1 **(138)** represent 5 x 5 sectors that are populated with zeros and so they are represented by blank squares with no circles.

The difference in position order of the sectors with nonzero values PDS1-IS1 **(136)** and PDS2-IS2 **(139)** could represent movement of the object or movement of the device capturing the image. The object features, in this case presentation of a unique set of pixel values in the sector, can be readily assessed by further processing of PDS1-I and PDS2-I by

MOPA **(3)**.

As illustrated in FIG. 1 and FIG. 3 and described in FIG. 6 Step **(31)**, PDS1 **(9)** and PDS2 **(10)** are input to MOPA **(3)** along with DEM **(8)** instruction set DEMM **(21)** and MC **(13)**, if any. Since this is the first pass through the NDDPP **(1)**, there is no MC **(13)** to input initially as described in FIG. 6 Step **(31)**. For the purposes of the example of the third embodiment of the present invention, DEMM-I will be used as a specific example of DEMM **(21)**.

FIG. 6, Step **(32)** describes the use of the integer values of PDS1 **(9)** and PDS2 **(10)** to make the matrix indices M1 **(11)** and M2 **(12)**. For the purposes of the example of the third embodiment of the present invention, M1-I will be used as a specific example of M1 **(11)** and M2-I will be used as a specific example of M2 **(12)**. As described in FIG. 6 Step **(32)**, the integer values in columns of the PDS1-I are used make M1-I indices. In this example of the third embodiment of the present invention, for M1-I: the value of the first index indicates number of each sector, such sector numbered sequentially in PDS1-I, the value of the second index represents the integer value of the pixel found in the sector, and the third line indicates the number of times that value was seen in that sector:

```
M1-I:

M1-I(  1,  0, 21 )
M1-I(  1,  2,  2 )
M1-I(  1,  6,  1 )
M1-I(  1,  8,  1 )
M1-I(  2,  0, 25 )
```

Similarly for M2-I: the value of the first index indicates number of each sector, such sector numbered sequentially in PDS2-I, the value of the second index represents the integer value of the pixel found in the sector, and the third line indicates the number of times that value was seen in that sector:

```
M2-I:
```

```
M2-I(  1,  0, 25 )
M2-I(  2,  0, 21 )
M2-I(  2,  2,  2 )
M2-I(  2,  6,  1 )
M2-I(  2,  8,  1 )
```

M1-I and M2-I are output by MOPA-MM **(22)** to USER **(5)** and also input to MOPA-CM **(23)** along with DEMM **(21)** as illustrated in FIG. 3 and described in FIG. 6 Steps **(32)** and **(33)**. There is no MC **(13)** input to the MOPA-CM **(23)** at this point in the example of the first embodiment, since MC **(13)** is an output of NDDPP **(1)** and it is the first pass of data through NDDPP **(1)**.

MOPA-CM **(23)** makes MC **(13)** as instructed by DEMM **(21)** as described in FIG. 6 Step **(33)**. DEMM **(21)** can instruct MOPA-CM **(23)** to output any combination of information related to shared indices of M1 **(11)** and M2 **(12)** and whether and how to use previously produced MC **(13)** in new MC **(13)** production. For the purposes of this example, MC-I will be used as an example of MC **(13)**. In this example, DEMM-I is an example of DEMM **(21)** and instructs MOPA-CM **(23)** to make MC-I.

```
DEMM-I:

FOR ALL M1-I AND M2-I INDICES:
  IF THE M1-I SECOND INDEX VALUE AND THE
    M2-I SECOND INDEX VALUE ARE THE SAME AND NOT ZERO,
  THEN MAKE AN NEW SET OF INDICES FOR MC-I:
    THE VALUE OF MC-I INDEX ONE GETS
      THE VALUE OF THE FIRST INDEX OF M1-I;
    THE VALUE OF MC-I INDEX TWO GETS
      THE VALUE OF THE FIRST INDEX OF M2-I; AND
    THE VALUE OF MC-I INDEX THREE GETS
      THE SHARED VALUE OF
        THE SECOND INDEX OF M1-I AND M2-I.
```

```
MC-I:

MC-I(  1,  2,  2 )
MC-I(  1,  2,  6 )
MC-I(  1,  2,  8 )
```

FIG. 19 is a view showing a visualization of MOPA **(3)** output of MC-I. FIG. 19 shows the three points whose x, y, and z coordinates are the values of the first, second, and third indices of MC-I. FIG. 17:

MC-I (1,2,2) is represented by a circle **(140)** at coordinate x=1, y=2, and z=2 indicating the pixel value "2" (third index of MC-I) occurring in the first sector of processed DS1-I, PDS1-IS1 (first index of MC-I) and in the second sector of processed DS2-I, PDS1-IS2 (second index of MC-I);

MC-I (1,2,6) is represented by a circle **(141)** at coordinate x=1, y=2, and z=6 indicating the pixel value "6" (third index of MC-I) occurring in the first sector of processed DS1-I, PDS1-IS1 (first index of MC-I) and in the second sector of processed DS2-I, PDS1-IS2 (second index of MC-I);

MC-I (1,2,8) is represented by a circle **(142)** at coordinate x=1, y=2, and z=8 indicating the pixel value "8" (third index of MC-I) occurring in the first sector of processed DS1-I, PDS1-IS1 (first index of MC-I) and in the second sector of processed DS2-I, PDS1-IS2 (second index of MC-I).

FIG. 19 is a graphical representation of MC-I and shows how MC-I can capture common spatial occupancy of pixels of a value in more than one image sector independent of orientation or arrangement of the common pixels and position of the sector in the data set or data sets.

In this example of the third embodiment, processed M1-I, M2-I, and MC-I are output

by MOPA **(3)** as illustrated in FIG. 3 and described in FIG. 6 Step **(34)**. M1-I, M2-I, and MC-I are output to USER **(5)**. MC-I is also output to FOPA **(4)** and MOPA-CM **(23)**.

FIG. 7 is a flowchart showing each step of the processing procedure of FOPA **(4)** and outlines how the example data is processed through FOPA **(4)** of the NDDPP **(1)** as illustrated in FIG. 1. FIG. 7 Step **(35)** describes the input of data sets DS1 **(6)**, DS2 **(7)** and MC **(13)** to FOPA-CC **(24)**. For the purposes of a detailed example of the processing of data by the NDDPP **(1)** of the present invention, the input data sets, DS1-I and DS2-I and MC-I, as described above, are input to FOPA-CC **(24)** as described in FIG. 7 Step **(35)**.

FOPA-CC **(24)** generates SCF **(25)** on the basis of MC-I as described in FIG.7 Step **(36)**. For the purposes of this example, SCF-I will be used as an example of the SCF **(25)**:

```
SCF-I:

SELECT ALL DS1-I INPUT SECTORS WHOSE SECTOR NUMBER IS THE
SAME AS MC-I INDEX ONE VALUE AND PUT ALL OF THE PIXEL
VALUES IN THE SECTOR IN DSC-I
```

In this example, as illustrated in FIG. 4, described in FIG. 7 Step **(37)**, and using SCF-I, FOPA-AC **(26)** outputs DSC-I, an example of DSC **(14)**:

```
 1.75   1.64   1.82    1.33   1.78
 1.20   1.19  20.09   20.14   1.56
 1.03  60.48   1.61    1.04   1.54
80.93   1.60   1.72    1.43   1.60
 1.24   1.58   1.70    1.06   1.60
```

DSC-I is output to DIPA-PA **(17)** and to USER **(5)** as illustrated in FIG. 1 and FIG. 4 and described in FIG. 7 Step **(38)**.

MC **(13)** output by MOPA-CM **(23)** is input back to MOPA-CM **(23)** and can provide feedback. Additional DEMM-I instructions, DEMM-Ia, can instruct MOPA-CM **(23)** to use MC **(13)** from a previous round of NDDPP **(1)** processing to instruct MOPA-CM **(23)** output

of a current round of processing. For example:

```
DEMM-Ia:

FOR ALL M1-I AND M2-I INDICES:
  IF THE M1-I SECOND INDEX VALUE, THE M2-I SECOND INDEX
   VALUE,
  AND THE MC-I THIRD INDEX ARE THE SAME,
    THEN MAKE AN NEW SET OF INDICES FOR MC-I:
      THE VALUE OF MC-I INDEX ONE GETS
        THE VALUE OF THE FIRST INDEX OF M1-I;
      THE VALUE OF MC-I INDEX TWO GETS
        THE VALUE OF THE FIRST INDEX OF M2-I; AND
      THE VALUE OF MC-I INDEX THIRD GETS
        THE SHARED VALUE OF
          THE SECOND INDEX OF M1-I AND M2-I.
```

DEMM-Ia ensures that common spatial occupancy of n-dimensional data elements from one round of NDDPP **(1)** processing are used in further rounds of NDDPP (**1)** processing. In this example, non-zero shared values are used to produce PDS1 **(9),** PDS2 **(10)**, M1 **(11),** M2 **(12)** , MC **(13)**, and DSC **(14)**. In this example, the DEMM-I instructs MOPA **(3)** that information about pixels that share a value of 0 are not to be reported. Conversely, if only values of zero that are shared are reported then PDS1 **(9)** and PDS2 **(10)** effectively become masks for data with pixels whose attributes can show a high degree of variability but can still be identified by exclusion of the most common pixel values. Data populated with common elements allow the tracking of pixels that may have highly variable attributes but are surrounded by pixels that have common elements, in this case those pixel values with a "0".

One skilled in the art would conclude that NDDPP **(1)** as described in this example can analyze any number of data sets with any number of dimensions or pixel attributes. In this example, information that may be image background is removed from consideration by eliminating certain shared values of pixels from consideration. 3-D pixels can be divided into

3-D sectors and processed as described in detail above using the present invention as described in detail in examples in the first and third embodiment.

Additional advantages and modifications will readily occur to those skilled in the art. Therefore, the invention in its broader aspects is not limited to the specific details and representative embodiments shown and described herein. Accordingly, various modifications may be made without departing from the spirit or scope of the general inventive concept as defined by the appended claims and their equivalents.

The foregoing descriptions of specific embodiments of the present invention have been presented for purposes of illustration and description. They are not intended to be exhaustive or to limit the invention to the precise forms disclosed, and obviously many modifications and variations are possible in light of the above teaching. The embodiments were chosen and described in order to best explain the principles of the invention and its practical application, to thereby enable others skilled in the art to best use the invention and various embodiments with various modifications as are suited to the particular use contemplated. It is intended that the scope of the invention be defined by the claims appended hereto and their equivalents.

TABLE 1

HETATM	4729	ZN	ZN	•	1	0.004	-0.434	-5.873	1.00	0.00	ZN

Table 1: DS1-M1 (the zinc (ZN) atom), excerpted from 1AIY.PDB, a specific example of DS1 **(6)**. ('•' substituted for space)

TABLE 2

ATOM	1	N	GLY	A	1	18.988	-5.893	-1.830	1.00	0.00	N
ATOM	2	CA	GLY	A	1	19.470	-4.536	-2.221	1.00	0.00	C
ATOM	3	C	GLY	A	1	18.271	-3.634	-2.517	1.00	0.00	C
ATOM	4	O	GLY	A	1	17.900	-2.795	-1.719	1.00	0.00	O
ATOM	10	N	ILE	A	2	17.663	-3.797	-3.659	1.00	0.00	N
ATOM	11	CA	ILE	A	2	16.492	-2.947	-4.008	1.00	0.00	C
ATOM	12	C	ILE	A	2	15.230	-3.509	-3.343	1.00	0.00	C
ATOM	13	O	ILE	A	2	14.210	-2.854	-3.275	1.00	0.00	O
ATOM	14	CB	ILE	A	2	16.316	-2.935	-5.527	1.00	0.00	C
ATOM	15	CG1	ILE	A	2	17.614	-2.455	-6.178	1.00	0.00	C
ATOM	16	CG2	ILE	A	2	15.175	-1.988	-5.901	1.00	0.00	C
ATOM	17	CD1	ILE	A	2	17.628	-2.854	-7.654	1.00	0.00	C

ATOM	29	N	VAL	A	3	15.294	-4.717	-2.851	1.00	0.00	N
ATOM	30	CA	VAL	A	3	14.103	-5.319	-2.190	1.00	0.00	C
ATOM	31	C	VAL	A	3	14.124	-4.992	-0.694	1.00	0.00	C
ATOM	32	O	VAL	A	3	13.603	-5.731	0.119	1.00	0.00	O
ATOM	33	CB	VAL	A	3	14.125	-6.838	-2.377	1.00	0.00	C
ATOM	34	CG1	VAL	A	3	12.828	-7.438	-1.832	1.00	0.00	C
ATOM	35	CG2	VAL	A	3	14.252	-7.164	-3.867	1.00	0.00	C
ATOM	45	N	GLU	A	4	14.722	-3.891	-0.324	1.00	0.00	N
ATOM	46	CA	GLU	A	4	14.774	-3.522	1.120	1.00	0.00	C
ATOM	47	C	GLU	A	4	15.261	-2.079	1.265	1.00	0.00	C
ATOM	48	O	GLU	A	4	14.660	-1.277	1.951	1.00	0.00	O
ATOM	49	CB	GLU	A	4	15.737	-4.456	1.858	1.00	0.00	C
ATOM	50	CG	GLU	A	4	15.335	-4.539	3.333	1.00	0.00	C
ATOM	51	CD	GLU	A	4	16.173	-5.606	4.037	1.00	0.00	C
ATOM	52	OE1	GLU	A	4	16.671	-6.487	3.355	1.00	0.00	O
ATOM	53	OE2	GLU	A	4	16.303	-5.524	5.248	1.00	0.00	O
ATOM	60	N	GLN	A	5	16.351	-1.742	0.629	1.00	0.00	N
ATOM	61	CA	GLN	A	5	16.878	-0.354	0.736	1.00	0.00	C
ATOM	62	C	GLN	A	5	15.892	0.620	0.088	1.00	0.00	C
ATOM	63	O	GLN	A	5	15.834	1.782	0.436	1.00	0.00	O
ATOM	64	CB	GLN	A	5	18.230	-0.263	0.026	1.00	0.00	C
ATOM	65	CG	GLN	A	5	18.733	1.180	0.073	1.00	0.00	C
ATOM	66	CD	GLN	A	5	19.907	1.347	-0.895	1.00	0.00	C
ATOM	67	OE1	GLN	A	5	21.030	1.026	-0.565	1.00	0.00	O
ATOM	68	NE2	GLN	A	5	19.692	1.839	-2.085	1.00	0.00	N
ATOM	77	N	CYS	A	6	15.114	0.153	-0.849	1.00	0.00	N
ATOM	78	CA	CYS	A	6	14.130	1.054	-1.513	1.00	0.00	C
ATOM	79	C	CYS	A	6	12.718	0.513	-1.289	1.00	0.00	C
ATOM	80	O	CYS	A	6	11.737	1.189	-1.528	1.00	0.00	O
ATOM	81	CB	CYS	A	6	14.418	1.118	-3.014	1.00	0.00	C
ATOM	82	SG	CYS	A	6	15.881	2.144	-3.307	1.00	0.00	S
ATOM	87	N	CYS	A	7	12.604	-0.702	-0.825	1.00	0.00	N
ATOM	88	CA	CYS	A	7	11.254	-1.281	-0.582	1.00	0.00	C
ATOM	89	C	CYS	A	7	10.742	-0.802	0.777	1.00	0.00	C
ATOM	90	O	CYS	A	7	9.563	-0.854	1.063	1.00	0.00	O
ATOM	91	CB	CYS	A	7	11.344	-2.809	-0.591	1.00	0.00	C
ATOM	92	SG	CYS	A	7	10.187	-3.479	-1.811	1.00	0.00	S
ATOM	97	N	THR	A	8	11.624	-0.331	1.615	1.00	0.00	N
ATOM	98	CA	THR	A	8	11.194	0.157	2.955	1.00	0.00	C
ATOM	99	C	THR	A	8	11.304	1.682	2.997	1.00	0.00	C
ATOM	100	O	THR	A	8	11.118	2.298	4.028	1.00	0.00	O
ATOM	101	CB	THR	A	8	12.095	-0.449	4.033	1.00	0.00	C
ATOM	102	OG1	THR	A	8	13.380	0.155	3.970	1.00	0.00	O
ATOM	103	CG2	THR	A	8	12.225	-1.958	3.807	1.00	0.00	C
ATOM	111	N	SER	A	9	11.606	2.297	1.886	1.00	0.00	N

ATOM	112	CA	SER	A	9	11.729	3.780	1.866	1.00	0.00	C
ATOM	113	C	SER	A	9	12.033	4.250	0.443	1.00	0.00	C
ATOM	114	O	SER	A	9	12.954	3.780	-0.194	1.00	0.00	O
ATOM	115	CB	SER	A	9	12.865	4.207	2.795	1.00	0.00	C
ATOM	116	OG	SER	A	9	14.007	3.400	2.541	1.00	0.00	O
ATOM	122	N	ILE	A	10	11.266	5.177	-0.061	1.00	0.00	N
ATOM	123	CA	ILE	A	10	11.514	5.675	-1.442	1.00	0.00	C
ATOM	124	C	ILE	A	10	12.980	6.083	-1.583	1.00	0.00	C
ATOM	125	O	ILE	A	10	13.626	6.460	-0.626	1.00	0.00	O
ATOM	126	CB	ILE	A	10	10.616	6.882	-1.713	1.00	0.00	C
ATOM	127	CG1	ILE	A	10	9.176	6.404	-1.900	1.00	0.00	C
ATOM	128	CG2	ILE	A	10	11.085	7.599	-2.982	1.00	0.00	C
ATOM	129	CD1	ILE	A	10	8.376	6.685	-0.628	1.00	0.00	C
ATOM	141	N	CYS	A	11	13.509	6.010	-2.773	1.00	0.00	N
ATOM	142	CA	CYS	A	11	14.932	6.392	-2.982	1.00	0.00	C
ATOM	143	C	CYS	A	11	15.011	7.499	-4.037	1.00	0.00	C
ATOM	144	O	CYS	A	11	14.008	8.054	-4.441	1.00	0.00	O
ATOM	145	CB	CYS	A	11	15.715	5.166	-3.458	1.00	0.00	C
ATOM	146	SG	CYS	A	11	15.530	3.830	-2.250	1.00	0.00	S
ATOM	151	N	SER	A	12	16.191	7.827	-4.488	1.00	0.00	N
ATOM	152	CA	SER	A	12	16.319	8.899	-5.515	1.00	0.00	C
ATOM	153	C	SER	A	12	16.695	8.277	-6.860	1.00	0.00	C
ATOM	154	O	SER	A	12	17.125	7.145	-6.934	1.00	0.00	O
ATOM	155	CB	SER	A	12	17.407	9.887	-5.094	1.00	0.00	C
ATOM	156	OG	SER	A	12	17.461	9.949	-3.675	1.00	0.00	O
ATOM	162	N	LEU	A	13	16.536	9.015	-7.922	1.00	0.00	N
ATOM	163	CA	LEU	A	13	16.884	8.477	-9.264	1.00	0.00	C
ATOM	164	C	LEU	A	13	18.385	8.180	-9.322	1.00	0.00	C
ATOM	165	O	LEU	A	13	18.822	7.266	-9.991	1.00	0.00	O
ATOM	166	CB	LEU	A	13	16.529	9.516	-10.328	1.00	0.00	C
ATOM	167	CG	LEU	A	13	16.555	8.864	-11.707	1.00	0.00	C
ATOM	168	CD1	LEU	A	13	15.202	9.057	-12.391	1.00	0.00	C
ATOM	169	CD2	LEU	A	13	17.651	9.516	-12.551	1.00	0.00	C
ATOM	181	N	TYR	A	14	19.175	8.950	-8.623	1.00	0.00	N
ATOM	182	CA	TYR	A	14	20.647	8.718	-8.636	1.00	0.00	C
ATOM	183	C	TYR	A	14	20.956	7.345	-8.037	1.00	0.00	C
ATOM	184	O	TYR	A	14	21.730	6.579	-8.577	1.00	0.00	O
ATOM	185	CB	TYR	A	14	21.341	9.797	-7.804	1.00	0.00	C
ATOM	186	CG	TYR	A	14	20.742	11.146	-8.116	1.00	0.00	C
ATOM	187	CD1	TYR	A	14	20.457	11.499	-9.441	1.00	0.00	C
ATOM	188	CD2	TYR	A	14	20.473	12.045	-7.078	1.00	0.00	C
ATOM	189	CE1	TYR	A	14	19.901	12.752	-9.726	1.00	0.00	C
ATOM	190	CE2	TYR	A	14	19.918	13.297	-7.363	1.00	0.00	C
ATOM	191	CZ	TYR	A	14	19.632	13.652	-8.686	1.00	0.00	C
ATOM	192	OH	TYR	A	14	19.083	14.886	-8.965	1.00	0.00	O

ATOM	202	N	GLN	A	15	20.364	7.032	-6.918	1.00	0.00	N
ATOM	203	CA	GLN	A	15	20.627	5.715	-6.277	1.00	0.00	C
ATOM	204	C	GLN	A	15	20.097	4.586	-7.168	1.00	0.00	C
ATOM	205	O	GLN	A	15	20.706	3.542	-7.287	1.00	0.00	O
ATOM	206	CB	GLN	A	15	19.928	5.666	-4.918	1.00	0.00	C
ATOM	207	CG	GLN	A	15	20.868	6.213	-3.842	1.00	0.00	C
ATOM	208	CD	GLN	A	15	20.078	6.468	-2.558	1.00	0.00	C
ATOM	209	OE1	GLN	A	15	18.994	7.014	-2.596	1.00	0.00	O
ATOM	210	NE2	GLN	A	15	20.579	6.094	-1.413	1.00	0.00	N
ATOM	219	N	LEU	A	16	18.968	4.786	-7.795	1.00	0.00	N
ATOM	220	CA	LEU	A	16	18.406	3.721	-8.676	1.00	0.00	C
ATOM	221	C	LEU	A	16	19.282	3.573	-9.923	1.00	0.00	C
ATOM	222	O	LEU	A	16	19.323	2.530	-10.546	1.00	0.00	O
ATOM	223	CB	LEU	A	16	16.986	4.104	-9.103	1.00	0.00	C
ATOM	224	CG	LEU	A	16	16.006	3.834	-7.960	1.00	0.00	C
ATOM	225	CD1	LEU	A	16	14.727	4.642	-8.188	1.00	0.00	C
ATOM	226	CD2	LEU	A	16	15.663	2.343	-7.921	1.00	0.00	C
ATOM	238	N	GLU	A	17	19.974	4.614	-10.298	1.00	0.00	N
ATOM	239	CA	GLU	A	17	20.840	4.545	-11.511	1.00	0.00	C
ATOM	240	C	GLU	A	17	22.063	3.669	-11.230	1.00	0.00	C
ATOM	241	O	GLU	A	17	22.678	3.139	-12.135	1.00	0.00	O
ATOM	242	CB	GLU	A	17	21.299	5.958	-11.884	1.00	0.00	C
ATOM	243	CG	GLU	A	17	21.696	6.000	-13.362	1.00	0.00	C
ATOM	244	CD	GLU	A	17	22.255	7.384	-13.698	1.00	0.00	C
ATOM	245	OE1	GLU	A	17	21.909	8.325	-13.004	1.00	0.00	O
ATOM	246	OE2	GLU	A	17	23.019	7.479	-14.644	1.00	0.00	O
ATOM	253	N	ASN	A	18	22.427	3.513	-9.988	1.00	0.00	N
ATOM	254	CA	ASN	A	18	23.614	2.674	-9.660	1.00	0.00	C
ATOM	255	C	ASN	A	18	23.298	1.200	-9.924	1.00	0.00	C
ATOM	256	O	ASN	A	18	24.185	0.373	-10.013	1.00	0.00	O
ATOM	257	CB	ASN	A	18	23.983	2.864	-8.187	1.00	0.00	C
ATOM	258	CG	ASN	A	18	24.426	4.310	-7.958	1.00	0.00	C
ATOM	259	OD1	ASN	A	18	24.926	4.954	-8.859	1.00	0.00	O
ATOM	260	ND2	ASN	A	18	24.262	4.851	-6.783	1.00	0.00	N
ATOM	267	N	TYR	A	19	22.044	0.862	-10.056	1.00	0.00	N
ATOM	268	CA	TYR	A	19	21.681	-0.559	-10.320	1.00	0.00	C
ATOM	269	C	TYR	A	19	21.565	-0.783	-11.830	1.00	0.00	C
ATOM	270	O	TYR	A	19	21.269	-1.871	-12.285	1.00	0.00	O
ATOM	271	CB	TYR	A	19	20.341	-0.879	-9.653	1.00	0.00	C
ATOM	272	CG	TYR	A	19	20.453	-0.673	-8.162	1.00	0.00	C
ATOM	273	CD1	TYR	A	19	21.494	-1.280	-7.448	1.00	0.00	C
ATOM	274	CD2	TYR	A	19	19.517	0.123	-7.492	1.00	0.00	C
ATOM	275	CE1	TYR	A	19	21.597	-1.091	-6.066	1.00	0.00	C
ATOM	276	CE2	TYR	A	19	19.620	0.312	-6.109	1.00	0.00	C
ATOM	277	CZ	TYR	A	19	20.661	-0.295	-5.396	1.00	0.00	C

ATOM	278	OH	TYR	A	19	20.762	-0.107	-4.033	1.00	0.00	O
ATOM	288	N	CYS	A	20	21.796	0.238	-12.614	1.00	0.00	N
ATOM	289	CA	CYS	A	20	21.697	0.079	-14.093	1.00	0.00	C
ATOM	290	C	CYS	A	20	22.814	-0.844	-14.585	1.00	0.00	C
ATOM	291	O	CYS	A	20	23.847	-0.973	-13.958	1.00	0.00	O
ATOM	292	CB	CYS	A	20	21.843	1.447	-14.766	1.00	0.00	C
ATOM	293	SG	CYS	A	20	20.209	2.200	-14.996	1.00	0.00	S
ATOM	298	N	ASN	A	21	22.614	-1.485	-15.703	1.00	0.00	N
ATOM	299	CA	ASN	A	21	23.663	-2.399	-16.239	1.00	0.00	C
ATOM	300	C	ASN	A	21	24.907	-1.588	-16.607	1.00	0.00	C
ATOM	301	O	ASN	A	21	25.947	-1.837	-16.023	1.00	0.00	O
ATOM	302	CB	ASN	A	21	23.132	-3.113	-17.484	1.00	0.00	C
ATOM	303	CG	ASN	A	21	22.914	-4.595	-17.172	1.00	0.00	C
ATOM	304	OD1	ASN	A	21	21.793	-5.061	-17.141	1.00	0.00	O
ATOM	305	ND2	ASN	A	21	23.945	-5.361	-16.938	1.00	0.00	N
ATOM	306	OXT	ASN	A	21	24.798	-0.731	-17.470	1.00	0.00	O

Table 2: DS2a-M1, the insulin A chain with hydrogens and waters removed, excerpted from 1AIY.PDB, a specific example of a DS2 **(7)**.

TABLE 3

ATOM	314	N	PHE	B	1	0.503	-8.361	6.002	1.00	0.00	N
ATOM	315	CA	PHE	B	1	0.748	-9.622	5.248	1.00	0.00	C
ATOM	316	C	PHE	B	1	1.312	-9.288	3.867	1.00	0.00	C
ATOM	317	O	PHE	B	1	2.509	-9.276	3.658	1.00	0.00	O
ATOM	318	CB	PHE	B	1	-0.567	-10.384	5.082	1.00	0.00	C
ATOM	319	CG	PHE	B	1	-0.775	-11.302	6.261	1.00	0.00	C
ATOM	320	CD1	PHE	B	1	-0.101	-12.526	6.318	1.00	0.00	C
ATOM	321	CD2	PHE	B	1	-1.643	-10.930	7.295	1.00	0.00	C
ATOM	322	CE1	PHE	B	1	-0.292	-13.380	7.409	1.00	0.00	C
ATOM	323	CE2	PHE	B	1	-1.836	-11.784	8.386	1.00	0.00	C
ATOM	324	CZ	PHE	B	1	-1.161	-13.010	8.444	1.00	0.00	C
ATOM	336	N	VAL	B	2	0.454	-9.019	2.922	1.00	0.00	N
ATOM	337	CA	VAL	B	2	0.929	-8.689	1.549	1.00	0.00	C
ATOM	338	C	VAL	B	2	1.634	-7.330	1.566	1.00	0.00	C
ATOM	339	O	VAL	B	2	1.269	-6.440	2.308	1.00	0.00	O
ATOM	340	CB	VAL	B	2	-0.270	-8.643	0.598	1.00	0.00	C
ATOM	341	CG1	VAL	B	2	-1.090	-9.925	0.757	1.00	0.00	C
ATOM	342	CG2	VAL	B	2	-1.150	-7.436	0.934	1.00	0.00	C
ATOM	352	N	ASN	B	3	2.648	-7.165	0.759	1.00	0.00	N
ATOM	353	CA	ASN	B	3	3.378	-5.866	0.739	1.00	0.00	C
ATOM	354	C	ASN	B	3	4.692	-6.033	-0.025	1.00	0.00	C
ATOM	355	O	ASN	B	3	5.165	-5.120	-0.673	1.00	0.00	O
ATOM	356	CB	ASN	B	3	3.680	-5.435	2.174	1.00	0.00	C

ATOM	357	CG	ASN	B	3	2.878	-4.178	2.512	1.00	0.00	C
ATOM	358	OD1	ASN	B	3	2.403	-3.491	1.631	1.00	0.00	O
ATOM	359	ND2	ASN	B	3	2.706	-3.848	3.763	1.00	0.00	N
ATOM	366	N	GLN	B	4	5.288	-7.192	0.050	1.00	0.00	N
ATOM	367	CA	GLN	B	4	6.568	-7.421	-0.669	1.00	0.00	C
ATOM	368	C	GLN	B	4	6.271	-7.749	-2.132	1.00	0.00	C
ATOM	369	O	GLN	B	4	7.020	-7.400	-3.022	1.00	0.00	O
ATOM	370	CB	GLN	B	4	7.316	-8.591	-0.026	1.00	0.00	C
ATOM	371	CG	GLN	B	4	8.612	-8.852	-0.798	1.00	0.00	C
ATOM	372	CD	GLN	B	4	9.718	-9.254	0.180	1.00	0.00	C
ATOM	373	OE1	GLN	B	4	10.620	-8.483	0.442	1.00	0.00	O
ATOM	374	NE2	GLN	B	4	9.686	-10.435	0.732	1.00	0.00	N
ATOM	383	N	HIS	B	5	5.178	-8.417	-2.387	1.00	0.00	N
ATOM	384	CA	HIS	B	5	4.833	-8.759	-3.793	1.00	0.00	C
ATOM	385	C	HIS	B	5	4.392	-7.488	-4.518	1.00	0.00	C
ATOM	386	O	HIS	B	5	4.507	-7.377	-5.722	1.00	0.00	O
ATOM	387	CB	HIS	B	5	3.695	-9.782	-3.812	1.00	0.00	C
ATOM	388	CG	HIS	B	5	3.603	-10.399	-5.183	1.00	0.00	C
ATOM	389	ND1	HIS	B	5	4.469	-11.397	-5.605	1.00	0.00	N
ATOM	390	CD2	HIS	B	5	2.759	-10.164	-6.240	1.00	0.00	C
ATOM	391	CE1	HIS	B	5	4.128	-11.723	-6.866	1.00	0.00	C
ATOM	392	NE2	HIS	B	5	3.093	-11.001	-7.301	1.00	0.00	N
ATOM	400	N	LEU	B	6	3.898	-6.523	-3.792	1.00	0.00	N
ATOM	401	CA	LEU	B	6	3.464	-5.254	-4.438	1.00	0.00	C
ATOM	402	C	LEU	B	6	4.594	-4.231	-4.332	1.00	0.00	C
ATOM	403	O	LEU	B	6	4.682	-3.301	-5.110	1.00	0.00	O
ATOM	404	CB	LEU	B	6	2.222	-4.711	-3.728	1.00	0.00	C
ATOM	405	CG	LEU	B	6	1.324	-5.869	-3.297	1.00	0.00	C
ATOM	406	CD1	LEU	B	6	0.091	-5.314	-2.579	1.00	0.00	C
ATOM	407	CD2	LEU	B	6	0.882	-6.658	-4.532	1.00	0.00	C
ATOM	419	N	CYS	B	7	5.460	-4.394	-3.369	1.00	0.00	N
ATOM	420	CA	CYS	B	7	6.584	-3.435	-3.203	1.00	0.00	C
ATOM	421	C	CYS	B	7	7.671	-3.733	-4.238	1.00	0.00	C
ATOM	422	O	CYS	B	7	8.422	-2.862	-4.631	1.00	0.00	O
ATOM	423	CB	CYS	B	7	7.166	-3.575	-1.795	1.00	0.00	C
ATOM	424	SG	CYS	B	7	8.508	-2.382	-1.569	1.00	0.00	S
ATOM	429	N	GLY	B	8	7.760	-4.956	-4.682	1.00	0.00	N
ATOM	430	CA	GLY	B	8	8.798	-5.308	-5.693	1.00	0.00	C
ATOM	431	C	GLY	B	8	8.253	-5.044	-7.097	1.00	0.00	C
ATOM	432	O	GLY	B	8	8.982	-5.057	-8.069	1.00	0.00	O
ATOM	436	N	SER	B	9	6.975	-4.803	-7.213	1.00	0.00	N
ATOM	437	CA	SER	B	9	6.383	-4.540	-8.556	1.00	0.00	C
ATOM	438	C	SER	B	9	6.350	-3.033	-8.816	1.00	0.00	C
ATOM	439	O	SER	B	9	6.233	-2.589	-9.942	1.00	0.00	O
ATOM	440	CB	SER	B	9	4.960	-5.095	-8.601	1.00	0.00	C

ATOM	441	OG	SER	B	9	4.256	-4.494	-9.680	1.00	0.00	O
ATOM	447	N	HIS	B	10	6.449	-2.243	-7.785	1.00	0.00	N
ATOM	448	CA	HIS	B	10	6.423	-0.768	-7.975	1.00	0.00	C
ATOM	449	C	HIS	B	10	7.833	-0.273	-8.299	1.00	0.00	C
ATOM	450	O	HIS	B	10	8.018	0.820	-8.796	1.00	0.00	O
ATOM	451	CB	HIS	B	10	5.929	-0.100	-6.692	1.00	0.00	C
ATOM	452	CG	HIS	B	10	4.435	0.059	-6.753	1.00	0.00	C
ATOM	453	ND1	HIS	B	10	3.819	0.907	-7.659	1.00	0.00	N
ATOM	454	CD2	HIS	B	10	3.421	-0.516	-6.027	1.00	0.00	C
ATOM	455	CE1	HIS	B	10	2.492	0.821	-7.457	1.00	0.00	C
ATOM	456	NE2	HIS	B	10	2.194	-0.032	-6.474	1.00	0.00	N
ATOM	464	N	LEU	B	11	8.832	-1.069	-8.025	1.00	0.00	N
ATOM	465	CA	LEU	B	11	10.227	-0.639	-8.321	1.00	0.00	C
ATOM	466	C	LEU	B	11	10.618	-1.108	-9.725	1.00	0.00	C
ATOM	467	O	LEU	B	11	11.480	-0.535	-10.362	1.00	0.00	O
ATOM	468	CB	LEU	B	11	11.189	-1.252	-7.300	1.00	0.00	C
ATOM	469	CG	LEU	B	11	10.623	-1.095	-5.888	1.00	0.00	C
ATOM	470	CD1	LEU	B	11	11.648	-1.601	-4.872	1.00	0.00	C
ATOM	471	CD2	LEU	B	11	10.326	0.382	-5.618	1.00	0.00	C
ATOM	483	N	VAL	B	12	9.993	-2.145	-10.214	1.00	0.00	N
ATOM	484	CA	VAL	B	12	10.338	-2.642	-11.577	1.00	0.00	C
ATOM	485	C	VAL	B	12	9.974	-1.571	-12.606	1.00	0.00	C
ATOM	486	O	VAL	B	12	10.718	-1.307	-13.532	1.00	0.00	O
ATOM	487	CB	VAL	B	12	9.557	-3.924	-11.877	1.00	0.00	C
ATOM	488	CG1	VAL	B	12	10.208	-4.654	-13.051	1.00	0.00	C
ATOM	489	CG2	VAL	B	12	9.571	-4.832	-10.645	1.00	0.00	C
ATOM	499	N	GLU	B	13	8.837	-0.948	-12.450	1.00	0.00	N
ATOM	500	CA	GLU	B	13	8.429	0.109	-13.416	1.00	0.00	C
ATOM	501	C	GLU	B	13	9.403	1.281	-13.318	1.00	0.00	C
ATOM	502	O	GLU	B	13	9.691	1.947	-14.293	1.00	0.00	O
ATOM	503	CB	GLU	B	13	7.021	0.602	-13.077	1.00	0.00	C
ATOM	504	CG	GLU	B	13	5.984	-0.385	-13.613	1.00	0.00	C
ATOM	505	CD	GLU	B	13	4.835	-0.508	-12.611	1.00	0.00	C
ATOM	506	OE1	GLU	B	13	3.962	0.344	-12.636	1.00	0.00	O
ATOM	507	OE2	GLU	B	13	4.849	-1.448	-11.835	1.00	0.00	O
ATOM	514	N	ALA	B	14	9.909	1.540	-12.146	1.00	0.00	N
ATOM	515	CA	ALA	B	14	10.858	2.667	-11.976	1.00	0.00	C
ATOM	516	C	ALA	B	14	12.293	2.153	-12.098	1.00	0.00	C
ATOM	517	O	ALA	B	14	13.231	2.804	-11.685	1.00	0.00	O
ATOM	518	CB	ALA	B	14	10.658	3.295	-10.597	1.00	0.00	C
ATOM	524	N	LEU	B	15	12.475	0.987	-12.659	1.00	0.00	N
ATOM	525	CA	LEU	B	15	13.852	0.439	-12.798	1.00	0.00	C
ATOM	526	C	LEU	B	15	14.281	0.494	-14.264	1.00	0.00	C
ATOM	527	O	LEU	B	15	15.265	1.117	-14.610	1.00	0.00	O
ATOM	528	CB	LEU	B	15	13.877	-1.014	-12.315	1.00	0.00	C

ATOM	529	CG	LEU	B	15	14.634	-1.104	-10.988	1.00	0.00	C
ATOM	530	CD1	LEU	B	15	14.474	-2.508	-10.405	1.00	0.00	C
ATOM	531	CD2	LEU	B	15	16.118	-0.819	-11.227	1.00	0.00	C
ATOM	543	N	TYR	B	16	13.555	-0.156	-15.130	1.00	0.00	N
ATOM	544	CA	TYR	B	16	13.931	-0.140	-16.570	1.00	0.00	C
ATOM	545	C	TYR	B	16	13.766	1.278	-17.128	1.00	0.00	C
ATOM	546	O	TYR	B	16	14.436	1.669	-18.062	1.00	0.00	O
ATOM	547	CB	TYR	B	16	13.034	-1.111	-17.346	1.00	0.00	C
ATOM	548	CG	TYR	B	16	11.656	-0.513	-17.514	1.00	0.00	C
ATOM	549	CD1	TYR	B	16	11.430	0.460	-18.494	1.00	0.00	C
ATOM	550	CD2	TYR	B	16	10.605	-0.933	-16.690	1.00	0.00	C
ATOM	551	CE1	TYR	B	16	10.156	1.016	-18.649	1.00	0.00	C
ATOM	552	CE2	TYR	B	16	9.328	-0.376	-16.845	1.00	0.00	C
ATOM	553	CZ	TYR	B	16	9.104	0.598	-17.825	1.00	0.00	C
ATOM	554	OH	TYR	B	16	7.848	1.148	-17.977	1.00	0.00	O
ATOM	564	N	LEU	B	17	12.879	2.052	-16.561	1.00	0.00	N
ATOM	565	CA	LEU	B	17	12.675	3.439	-17.061	1.00	0.00	C
ATOM	566	C	LEU	B	17	13.889	4.296	-16.700	1.00	0.00	C
ATOM	567	O	LEU	B	17	14.238	5.223	-17.404	1.00	0.00	O
ATOM	568	CB	LEU	B	17	11.421	4.037	-16.423	1.00	0.00	C
ATOM	569	CG	LEU	B	17	11.261	5.489	-16.874	1.00	0.00	C
ATOM	570	CD1	LEU	B	17	11.302	5.558	-18.402	1.00	0.00	C
ATOM	571	CD2	LEU	B	17	9.919	6.029	-16.375	1.00	0.00	C
ATOM	583	N	VAL	B	18	14.537	3.994	-15.607	1.00	0.00	N
ATOM	584	CA	VAL	B	18	15.730	4.792	-15.204	1.00	0.00	C
ATOM	585	C	VAL	B	18	16.870	4.540	-16.193	1.00	0.00	C
ATOM	586	O	VAL	B	18	17.384	5.454	-16.807	1.00	0.00	O
ATOM	587	CB	VAL	B	18	16.176	4.376	-13.802	1.00	0.00	C
ATOM	588	CG1	VAL	B	18	17.540	4.997	-13.495	1.00	0.00	C
ATOM	589	CG2	VAL	B	18	15.153	4.868	-12.777	1.00	0.00	C
ATOM	599	N	CYS	B	19	17.269	3.306	-16.355	1.00	0.00	N
ATOM	600	CA	CYS	B	19	18.374	3.001	-17.307	1.00	0.00	C
ATOM	601	C	CYS	B	19	18.043	3.605	-18.674	1.00	0.00	C
ATOM	602	O	CYS	B	19	18.830	4.328	-19.252	1.00	0.00	O
ATOM	603	CB	CYS	B	19	18.535	1.483	-17.445	1.00	0.00	C
ATOM	604	SG	CYS	B	19	19.098	0.760	-15.879	1.00	0.00	S
ATOM	609	N	GLY	B	20	16.883	3.314	-19.195	1.00	0.00	N
ATOM	610	CA	GLY	B	20	16.501	3.871	-20.524	1.00	0.00	C
ATOM	611	C	GLY	B	20	16.656	2.789	-21.598	1.00	0.00	C
ATOM	612	O	GLY	B	20	15.993	1.772	-21.565	1.00	0.00	O
ATOM	616	N	GLU	B	21	17.527	3.000	-22.550	1.00	0.00	N
ATOM	617	CA	GLU	B	21	17.722	1.983	-23.622	1.00	0.00	C
ATOM	618	C	GLU	B	21	18.485	0.782	-23.055	1.00	0.00	C
ATOM	619	O	GLU	B	21	18.608	-0.246	-23.691	1.00	0.00	O
ATOM	620	CB	GLU	B	21	18.526	2.596	-24.772	1.00	0.00	C

ATOM	621	CG	GLU	B	21	17.622	2.774	-25.994	1.00	0.00	C
ATOM	622	CD	GLU	B	21	18.060	4.015	-26.774	1.00	0.00	C
ATOM	623	OE1	GLU	B	21	19.195	4.428	-26.601	1.00	0.00	O
ATOM	624	OE2	GLU	B	21	17.253	4.531	-27.529	1.00	0.00	O
ATOM	631	N	ARG	B	22	19.002	0.908	-21.863	1.00	0.00	N
ATOM	632	CA	ARG	B	22	19.760	-0.220	-21.255	1.00	0.00	C
ATOM	633	C	ARG	B	22	18.800	-1.121	-20.474	1.00	0.00	C
ATOM	634	O	ARG	B	22	17.645	-0.792	-20.282	1.00	0.00	O
ATOM	635	CB	ARG	B	22	20.821	0.339	-20.305	1.00	0.00	C
ATOM	636	CG	ARG	B	22	22.202	0.218	-20.952	1.00	0.00	C
ATOM	637	CD	ARG	B	22	23.039	1.446	-20.598	1.00	0.00	C
ATOM	638	NE	ARG	B	22	24.440	1.240	-21.060	1.00	0.00	N
ATOM	639	CZ	ARG	B	22	25.194	2.267	-21.335	1.00	0.00	C
ATOM	640	NH1	ARG	B	22	24.942	2.997	-22.387	1.00	0.00	N
ATOM	641	NH2	ARG	B	22	26.202	2.562	-20.561	1.00	0.00	N
ATOM	655	N	GLY	B	23	19.264	-2.253	-20.022	1.00	0.00	N
ATOM	656	CA	GLY	B	23	18.372	-3.170	-19.257	1.00	0.00	C
ATOM	657	C	GLY	B	23	18.838	-3.247	-17.801	1.00	0.00	C
ATOM	658	O	GLY	B	23	19.565	-2.397	-17.326	1.00	0.00	O
ATOM	662	N	PHE	B	24	18.425	-4.261	-17.090	1.00	0.00	N
ATOM	663	CA	PHE	B	24	18.841	-4.396	-15.663	1.00	0.00	C
ATOM	664	C	PHE	B	24	18.456	-5.787	-15.152	1.00	0.00	C
ATOM	665	O	PHE	B	24	17.786	-6.541	-15.829	1.00	0.00	O
ATOM	666	CB	PHE	B	24	18.132	-3.329	-14.825	1.00	0.00	C
ATOM	667	CG	PHE	B	24	16.656	-3.639	-14.761	1.00	0.00	C
ATOM	668	CD1	PHE	B	24	15.886	-3.610	-15.930	1.00	0.00	C
ATOM	669	CD2	PHE	B	24	16.056	-3.959	-13.537	1.00	0.00	C
ATOM	670	CE1	PHE	B	24	14.516	-3.896	-15.875	1.00	0.00	C
ATOM	671	CE2	PHE	B	24	14.685	-4.246	-13.481	1.00	0.00	C
ATOM	672	CZ	PHE	B	24	13.915	-4.216	-14.651	1.00	0.00	C
ATOM	682	N	PHE	B	25	18.867	-6.134	-13.964	1.00	0.00	N
ATOM	683	CA	PHE	B	25	18.516	-7.477	-13.424	1.00	0.00	C
ATOM	684	C	PHE	B	25	17.718	-7.317	-12.128	1.00	0.00	C
ATOM	685	O	PHE	B	25	17.829	-6.325	-11.435	1.00	0.00	O
ATOM	686	CB	PHE	B	25	19.794	-8.270	-13.143	1.00	0.00	C
ATOM	687	CG	PHE	B	25	20.577	-7.595	-12.041	1.00	0.00	C
ATOM	688	CD1	PHE	B	25	21.393	-6.495	-12.333	1.00	0.00	C
ATOM	689	CD2	PHE	B	25	20.487	-8.071	-10.727	1.00	0.00	C
ATOM	690	CE1	PHE	B	25	22.120	-5.872	-11.311	1.00	0.00	C
ATOM	691	CE2	PHE	B	25	21.213	-7.449	-9.706	1.00	0.00	C
ATOM	692	CZ	PHE	B	25	22.030	-6.349	-9.998	1.00	0.00	C
ATOM	702	N	TYR	B	26	16.920	-8.294	-11.797	1.00	0.00	N
ATOM	703	CA	TYR	B	26	16.112	-8.219	-10.548	1.00	0.00	C
ATOM	704	C	TYR	B	26	16.088	-9.598	-9.886	1.00	0.00	C
ATOM	705	O	TYR	B	26	15.545	-10.544	-10.420	1.00	0.00	O

ATOM	706	CB	TYR	B	26	14.681	-7.793	-10.883	1.00	0.00	C
ATOM	707	CG	TYR	B	26	13.857	-7.763	-9.617	1.00	0.00	C
ATOM	708	CD1	TYR	B	26	14.305	-7.037	-8.505	1.00	0.00	C
ATOM	709	CD2	TYR	B	26	12.645	-8.460	-9.554	1.00	0.00	C
ATOM	710	CE1	TYR	B	26	13.540	-7.011	-7.331	1.00	0.00	C
ATOM	711	CE2	TYR	B	26	11.881	-8.434	-8.380	1.00	0.00	C
ATOM	712	CZ	TYR	B	26	12.327	-7.709	-7.269	1.00	0.00	C
ATOM	713	OH	TYR	B	26	11.573	-7.683	-6.112	1.00	0.00	O
ATOM	723	N	THR	B	27	16.673	-9.722	-8.730	1.00	0.00	N
ATOM	724	CA	THR	B	27	16.680	-11.042	-8.044	1.00	0.00	C
ATOM	725	C	THR	B	27	16.583	-10.826	-6.531	1.00	0.00	C
ATOM	726	O	THR	B	27	17.555	-10.466	-5.898	1.00	0.00	O
ATOM	727	CB	THR	B	27	17.979	-11.780	-8.373	1.00	0.00	C
ATOM	728	OG1	THR	B	27	18.409	-11.421	-9.680	1.00	0.00	O
ATOM	729	CG2	THR	B	27	17.737	-13.289	-8.308	1.00	0.00	C
ATOM	737	N	PRO	B	28	15.407	-11.046	-5.996	1.00	0.00	N
ATOM	738	CA	PRO	B	28	15.153	-10.876	-4.555	1.00	0.00	C
ATOM	739	C	PRO	B	28	15.714	-12.067	-3.773	1.00	0.00	C
ATOM	740	O	PRO	B	28	15.103	-13.114	-3.691	1.00	0.00	O
ATOM	741	CB	PRO	B	28	13.625	-10.824	-4.464	1.00	0.00	C
ATOM	742	CG	PRO	B	28	13.086	-11.518	-5.737	1.00	0.00	C
ATOM	743	CD	PRO	B	28	14.228	-11.484	-6.771	1.00	0.00	C
ATOM	751	N	LYS	B	29	16.875	-11.912	-3.196	1.00	0.00	N
ATOM	752	CA	LYS	B	29	17.473	-13.035	-2.419	1.00	0.00	C
ATOM	753	C	LYS	B	29	18.649	-12.519	-1.587	1.00	0.00	C
ATOM	754	O	LYS	B	29	19.239	-11.500	-1.889	1.00	0.00	O
ATOM	755	CB	LYS	B	29	17.967	-14.119	-3.380	1.00	0.00	C
ATOM	756	CG	LYS	B	29	18.108	-15.443	-2.625	1.00	0.00	C
ATOM	757	CD	LYS	B	29	18.033	-16.608	-3.615	1.00	0.00	C
ATOM	758	CE	LYS	B	29	18.086	-17.932	-2.850	1.00	0.00	C
ATOM	759	NZ	LYS	B	29	19.277	-18.712	-3.292	1.00	0.00	N
ATOM	773	N	THR	B	30	18.995	-13.218	-0.540	1.00	0.00	N
ATOM	774	CA	THR	B	30	20.131	-12.777	0.317	1.00	0.00	C
ATOM	775	C	THR	B	30	21.332	-13.696	0.077	1.00	0.00	C
ATOM	776	O	THR	B	30	22.432	-13.182	-0.037	1.00	0.00	O
ATOM	777	CB	THR	B	30	19.717	-12.852	1.790	1.00	0.00	C
ATOM	778	OG1	THR	B	30	18.309	-12.702	1.891	1.00	0.00	O
ATOM	779	CG2	THR	B	30	20.410	-11.739	2.581	1.00	0.00	C
ATOM	780	OXT	THR	B	30	21.129	-14.897	0.013	1.00	0.00	O

Table 3: DS2b-M1, the insulin B chain with hydrogens and waters removed, excerpted from 1AIY.PDB, a specific example of more than one input data set as DS2 **(7)**.

CLAIMS

What is claimed is:

1. An n-dimensional data pipeline processor comprising:

a data input preparing apparatus for:

inputting one or more n-dimensional data sets;

inputting data entry modifier instructions that direct the parsing, attributing, shifting, sizing, and truncating of input data;

parsing, attributing, shifting, sizing, and truncating data according to data entry modifier instructions; and

outputting parsed, attributed, shifted, sized, and truncated data; and

a matrix output preparing apparatus for:

inputting parsed, attributed, shifted, sized, and truncated data;

making matrix indices that represent the characterized spatial occupancy of n-dimensional data elements by using parsed, attributed, shifted, sized, and truncated data as the matrix index values;

outputting matrix indices that represent the characterized spatial occupancy;

inputting matrix indices made by the matrix output preparing apparatus;

inputting data entry modifier instructions that direct the production of matrix indices from matrix indices that have shared matrix index values;

making matrix indices that represent the common spatial occupancy of n-dimensional data elements using the data entry modifier instructions to direct the production of matrix indices from matrix indices that have shared matrix index values;

outputting matrix indices that represent the common spatial occupancy; and

a feature output preparing apparatus for:

inputting matrix indices that represent the common spatial occupancy;

making selection criteria based on matrix indices that represent the common spatial occupancy of n-dimensional data elements;

selecting data from n-dimensional data using selection criteria; and

outputting selected n-dimensional data.

2. The n-dimensional data pipeline processor of claim 1, wherein input n-dimensional data

and input data entry modifier instructions are input by USER.

3. The n-dimensional data pipeline processor of claim 1, wherein the output matrix indices and output selected data are output to USER.

4. The n-dimensional data pipeline processor of claim 1, wherein the output selected n-dimensional data made by the feature output preparing apparatus is input to the data input preparing apparatus.

5. The n-dimensional data pipeline processor of claim 1, wherein the matrix indices that represent common spatial occupancy of n-dimensional data elements are input to the matrix output preparing apparatus for use in producing additional matrix indices that represent common spatial occupancy of n-dimensional data elements.

6. The data entry modifier instructions according to claim 1, wherein the data entry modifier instructions incorporate information derived from the input n-dimensional data.

7. The data entry modifier instructions according to claim 6, wherein the data entry modifier instructions incorporate information derived from the values of input n-dimensional data.

8. The data entry modifier instructions according to claim 6, wherein the data entry modifier instructions incorporate information derived from the number of unique data elements in the input n-dimensional data.

9. The data entry modifier instructions that direct the parsing, attributing, shifting, sizing, and truncating of input data according to claim 1, wherein the data entry modifier instructions that direct the parsing, attributing, shifting, sizing, and truncating of input data include directions and values for producing multiple shifted data sets.

10. The data entry modifier instructions that direct the parsing, attributing, shifting, sizing, and truncating of input data according to claim 1, wherein the data entry modifier instructions that direct the parsing, attributing, shifting, sizing, and truncating of input data include directions and values for producing multiple sized data sets.

11. The selected n-dimensional data according to claim 1, wherein the selected n-dimensional data represents a feature of the input data set.

12. The method of making matrix indices comprising using parsed, attributed, shifted, sized, and truncated values of n-dimensional data as matrix index values.

13. The method of making matrix indices by using matrix index values from matrix indices that share common values in two or more matrix index positions and whose matrix index values are parsed, attributed, shifted, sized, and truncated values of n-dimensional data.

14. A method of processing n-dimensional data comprising:

inputting n-dimensional data;

parsing, attributing, shifting, sizing, and truncating input n-dimensional data;

making matrix indices that represent characterized spatial occupancy of n-dimensional data elements by using parsed, attributed, shifted, sized, and truncated data as the matrix index values;

making matrix indices that represent common spatial occupancy of n-dimensional data elements from matrix indices that have shared matrix index values that are derived from parsed, attributed, shifted, sized, and truncated n-dimensional data;

outputting matrix indices that represent the common spatial occupancy of n-dimensional data elements;

making selection criteria using processed matrix indices that represent common spatial occupancy of n-dimensional data elements;

using selection criteria to select input n-dimensional data; and

outputting selected n-dimensional data.

15. The method of processing n-dimensional data of claim 14, wherein matrix indices that represent common spatial occupancy of n-dimensional data elements are used to make additional matrix indices that represent common spatial occupancy of n-dimensional data elements.

16. The method of processing n-dimensional data of claim 14, wherein the output selected n-dimensional data is used as input n-dimensional data.

17. The method of processing n-dimensional data of claim 14, wherein the parsing, attributing, shifting, sizing, and truncating of input n-dimensional data and making matrix indices that represent common spatial occupancy of n-dimensional data elements are directed by data entry modifier instructions.

18. The method of using matrix indices that represent common spatial occupancy of n-

dimensional data elements to identify data elements that occupy a common spatial volume.

19. The method of using matrix indices that represent common spatial occupancy of n-dimensional data elements to identify the same data elements that occupy different spatial volumes of identical size.

20. The method of using matrix indices that represent common spatial occupancy of n-dimensional data elements to identify the same data elements that occupy overlapping spatial volumes of different size.

21. The method of using matrix indices that represent common spatial occupancy of n-dimensional data elements to identify the same data elements that occupy non-overlapping spatial volumes of different size.

22. The method of using common spatial occupancy of n-dimensional data elements to identify spatial volumes with common elements.

23. The method of systematically changing the value of a shift value used to shift a parsed and attributed data element to produce multiple sets of parsed, attributed, and shifted data in order to produce multiple sets of matrix indices that represent characterized spatial occupancy of n-dimensional data elements.

24. The method of sytematically changing the value of a multiply value used to size a parsed, attributed, and shifted data element to produce multiple sets of parsed, attributed, shifted, and sized data in order to produce multiple sets of matrix indices that represent characterized spatial occupancy of n-dimensional data elements.

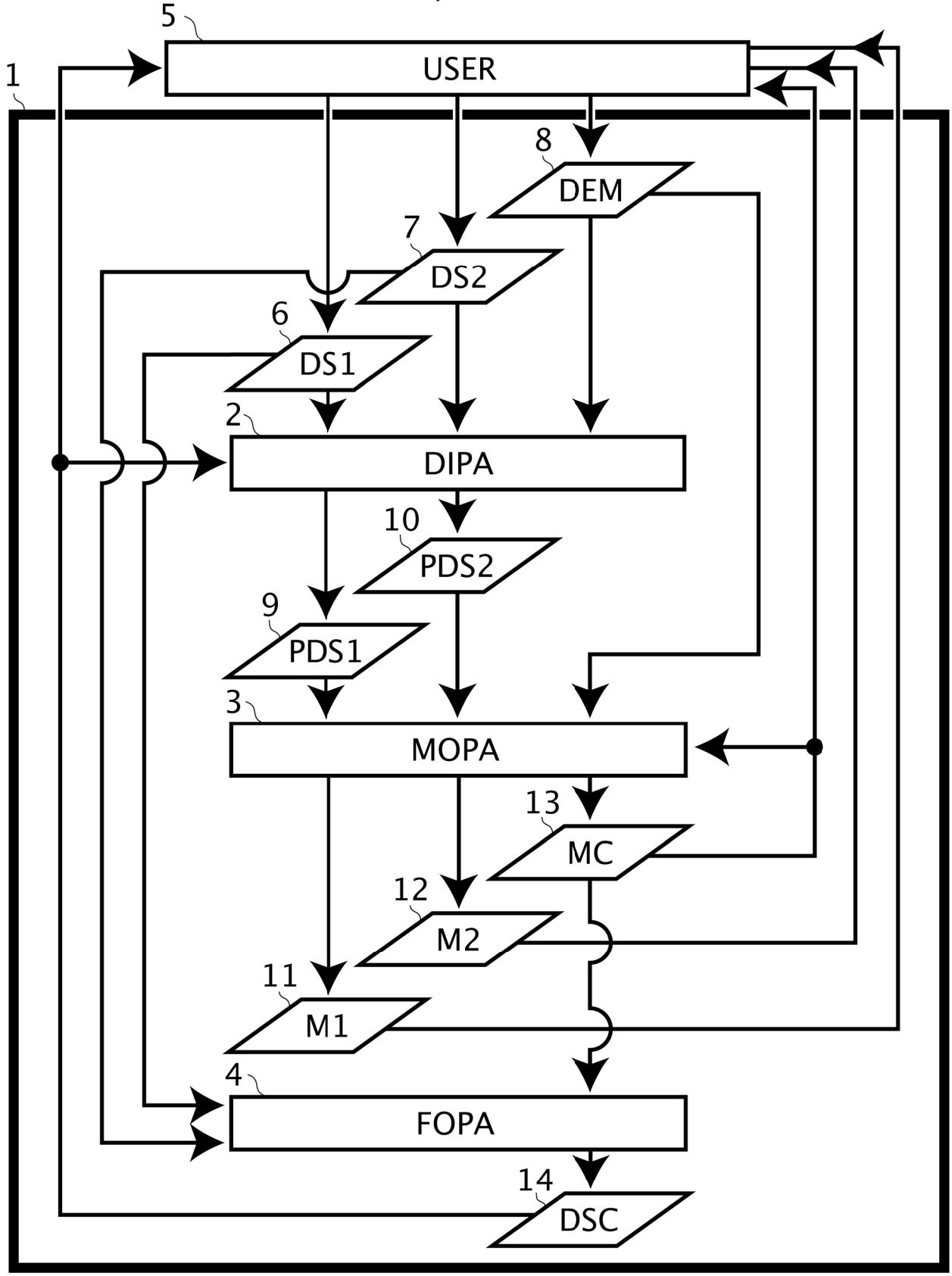
5
USER
1
8
DEM
7
DS2
6
DS1
2
DIPA
10
PDS2
9
PDS1
3
MOPA
13
MC
12
M2
11
M1
4
FOPA
14
DSC

FIGURE 1

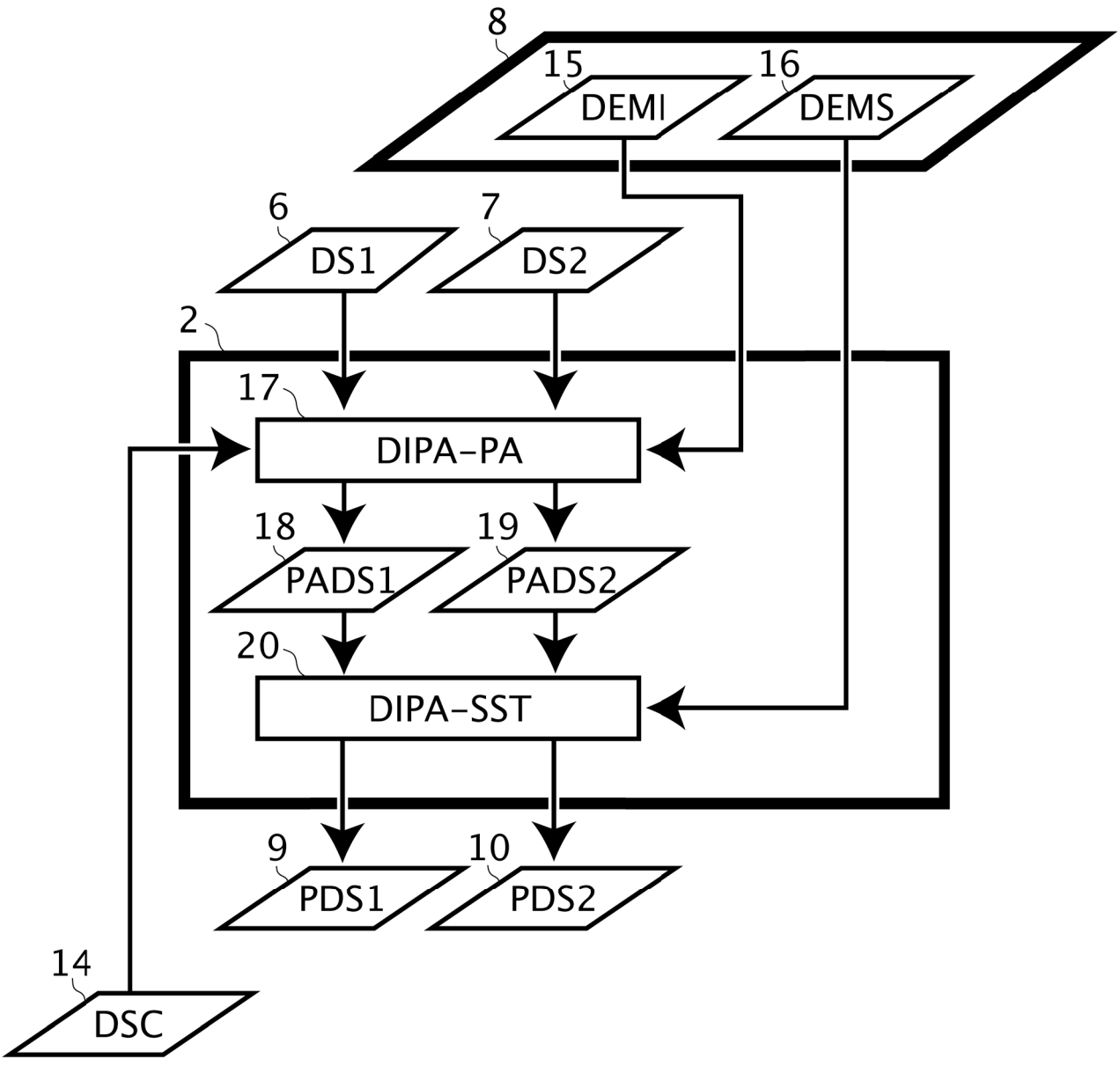
8
15
DEMI
16
DEMS
6
DS1
7
DS2
2
17
DIPA-PA
18
PADS1
19
PADS2
20
DIPA-SST
9
PDS1
10
PDS2
14
DSC

FIGURE 2

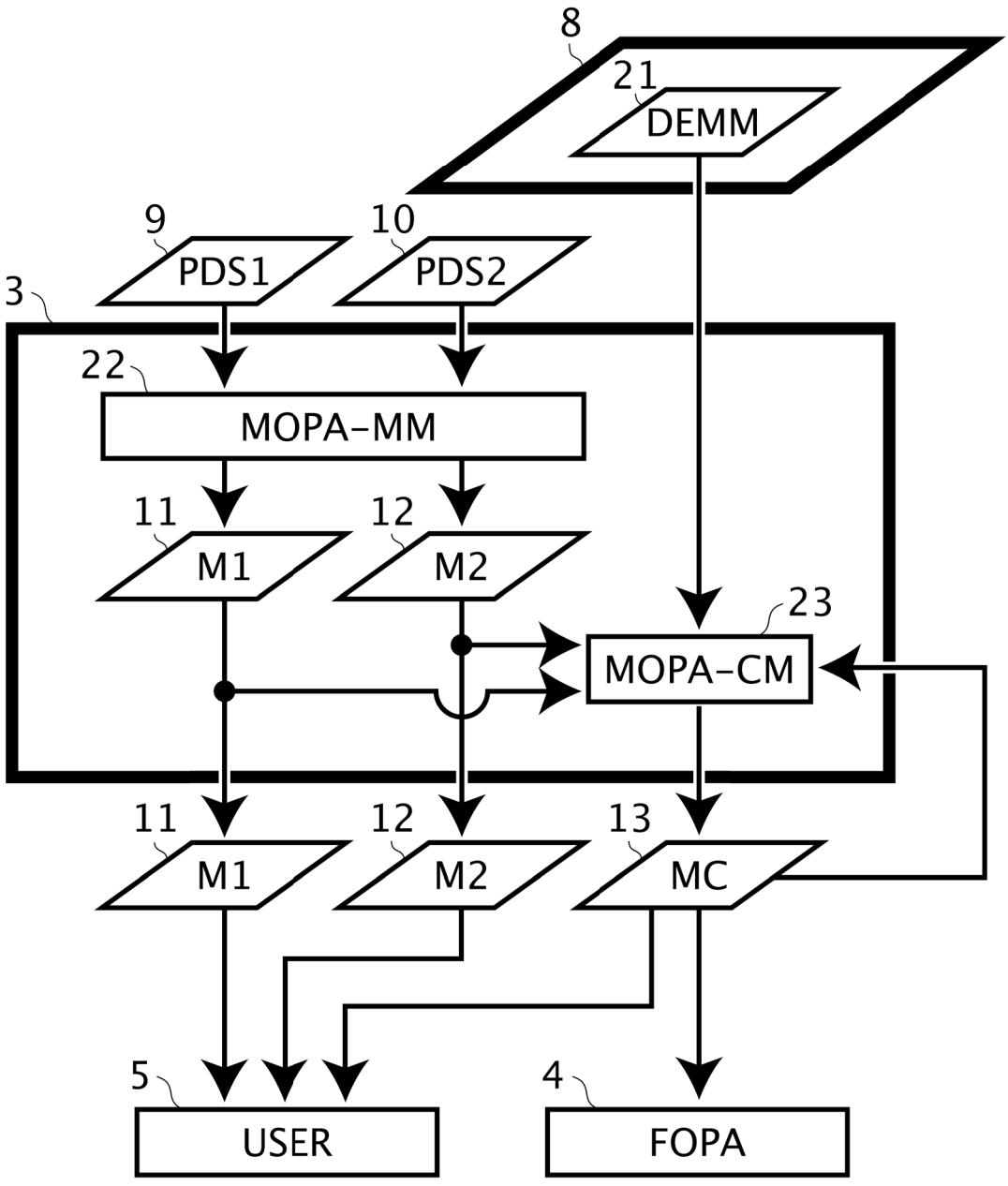
8
21
DEMM
9
PDS1
10
PDS2
3
22
MOPA-MM
11
M1
12
M2
23
MOPA-CM
11
M1
12
M2
13
MC
5
USER
4
FOPA

FIGURE 3

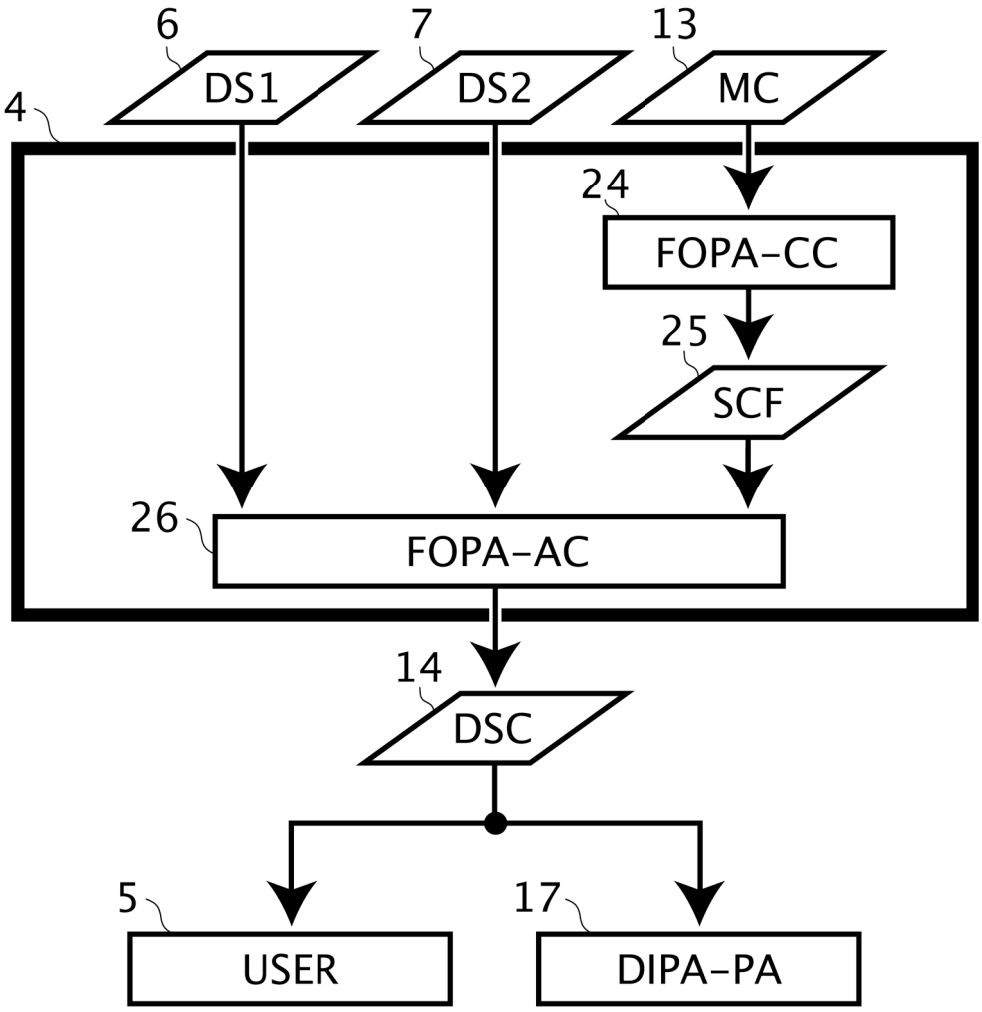
6
DS1
7
DS2
13
MC
4
24
FOPA-CC
25
SCF
26
FOPA-AC
14
DSC
5
USER
17
DIPA-PA

FIGURE 4

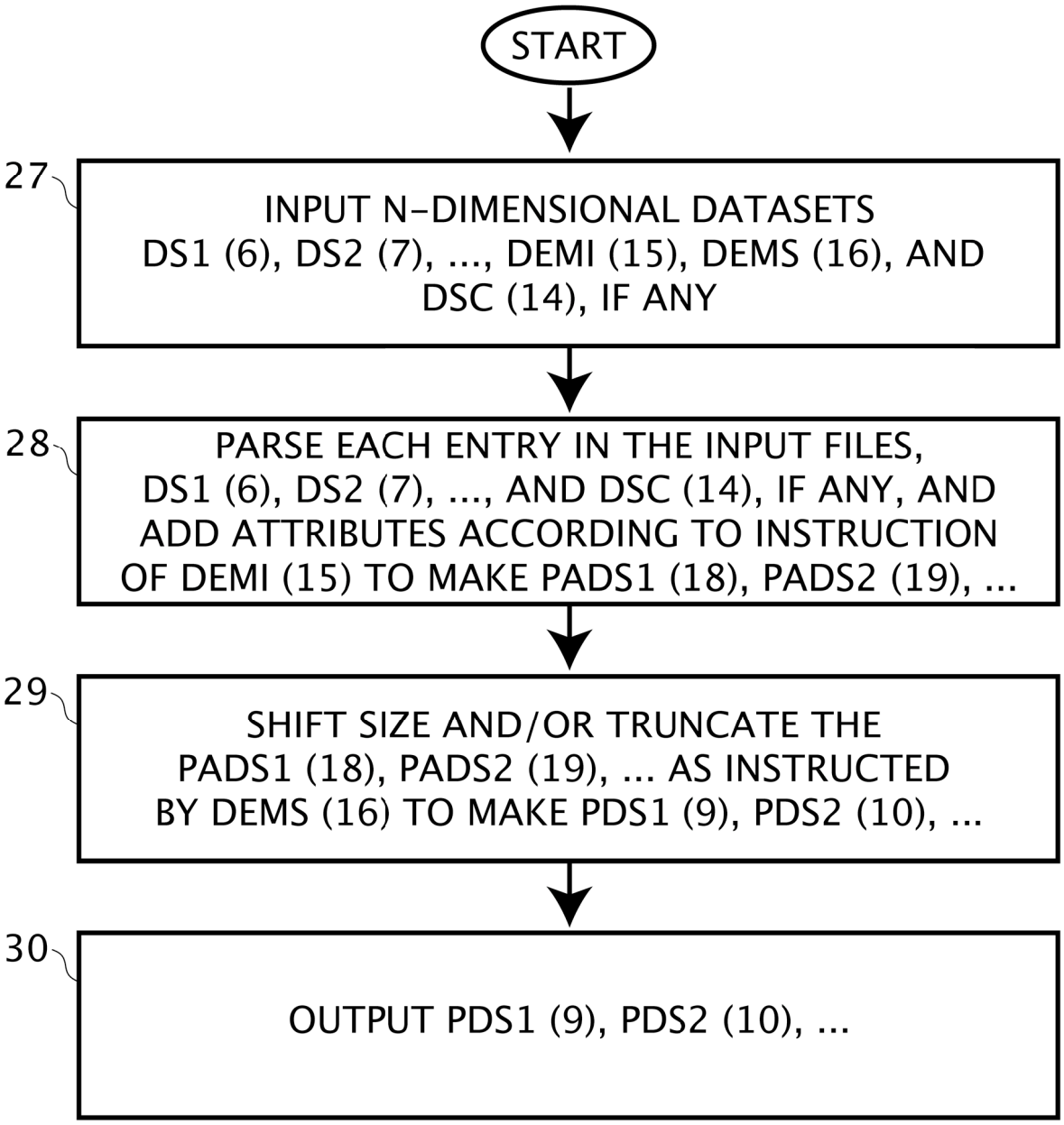
START
27
INPUT N-DIMENSIONAL DATASETS DS1 (6), DS2 (7), ..., DEMI (15), DEMS (16), AND DSC (14), IF ANY
28
PARSE EACH ENTRY IN THE INPUT FILES, DS1 (6), DS2 (7), ..., AND DSC (14), IF ANY, AND ADD ATTRIBUTES ACCORDING TO INSTRUCTION OF DEMI (15) TO MAKE PADS1 (18), PADS2 (19), ...
29
SHIFT SIZE AND/OR TRUNCATE THE PADS1 (18), PADS2 (19), ... AS INSTRUCTED BY DEMS (16) TO MAKE PDS1 (9), PDS2 (10), ...
30
OUTPUT PDS1 (9), PDS2 (10), ...

FIGURE 5

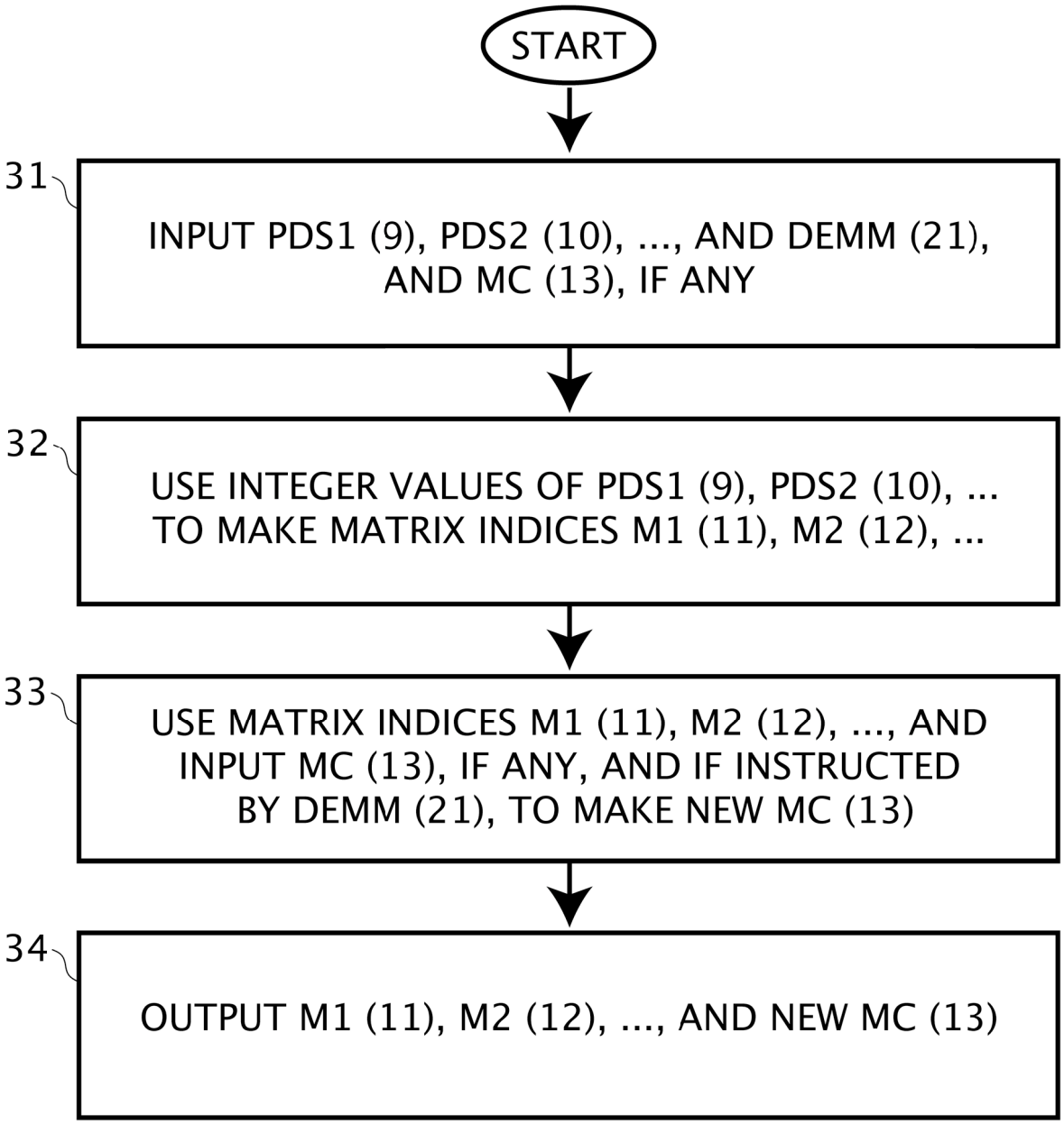
START
31
INPUT PDS1 (9), PDS2 (10), ..., AND DEMM (21), AND MC (13), IF ANY
32
USE INTEGER VALUES OF PDS1 (9), PDS2 (10), ... TO MAKE MATRIX INDICES M1 (11), M2 (12), ...
33
USE MATRIX INDICES M1 (11), M2 (12), ..., AND INPUT MC (13), IF ANY, AND IF INSTRUCTED BY DEMM (21), TO MAKE NEW MC (13)
34
OUTPUT M1 (11), M2 (12), ..., AND NEW MC (13)

FIGURE 6

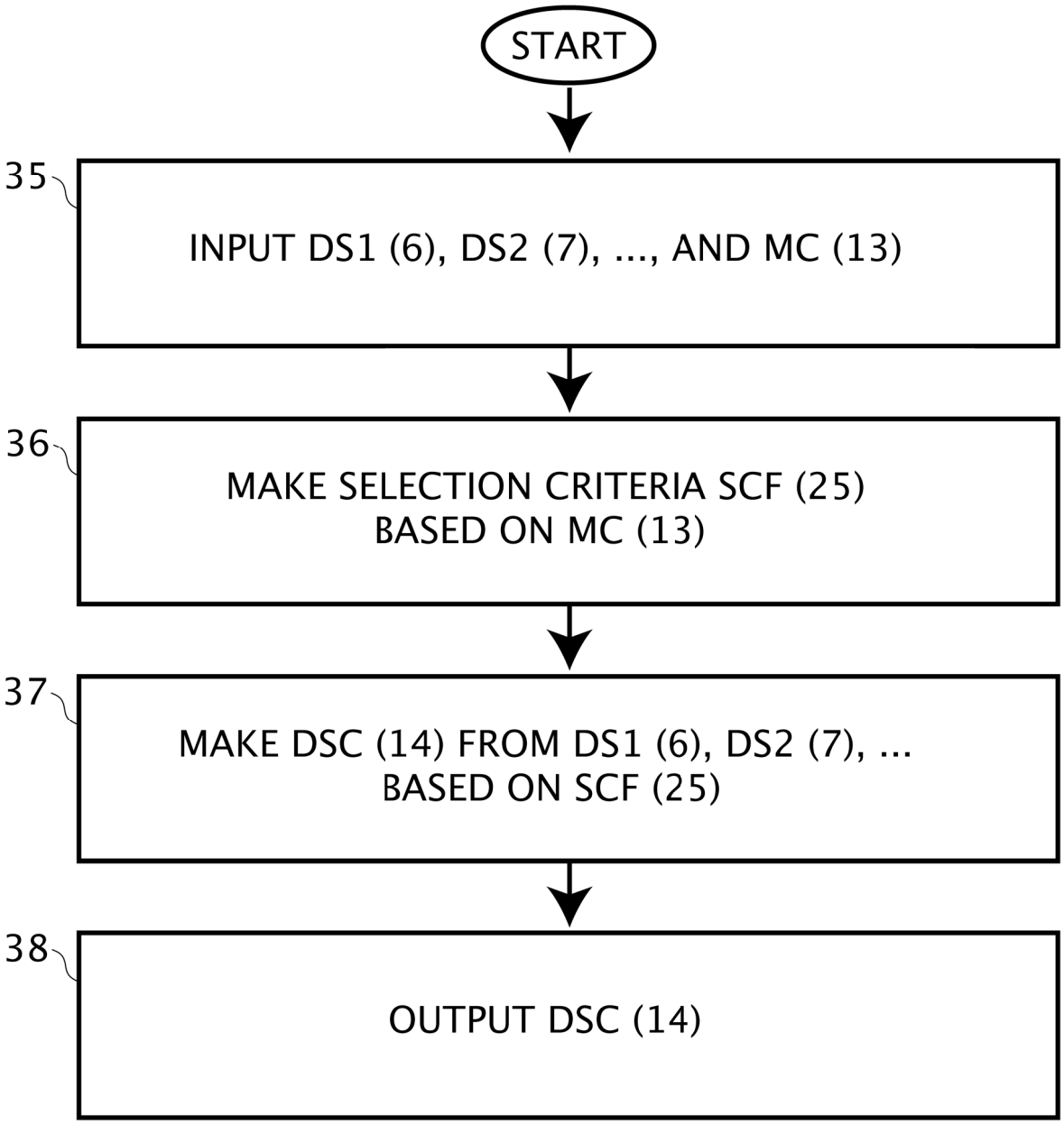
START
35
INPUT DS1 (6), DS2 (7), ..., AND MC (13)
36
MAKE SELECTION CRITERIA SCF (25)
BASED ON MC (13)
37
MAKE DSC (14) FROM DS1 (6), DS2 (7), ...
BASED ON SCF (25)
38
OUTPUT DSC (14)

FIGURE 7

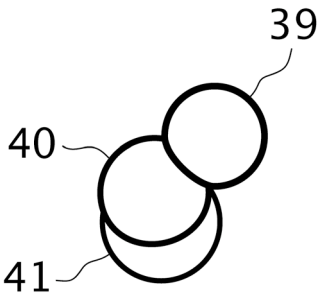

FIGURE 8

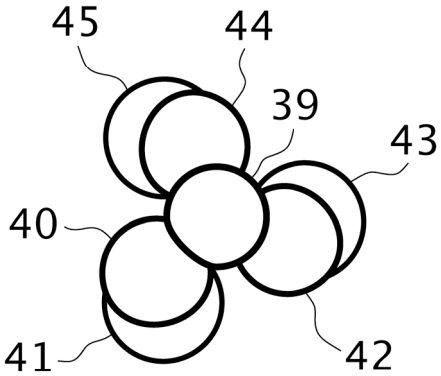

FIGURE 9

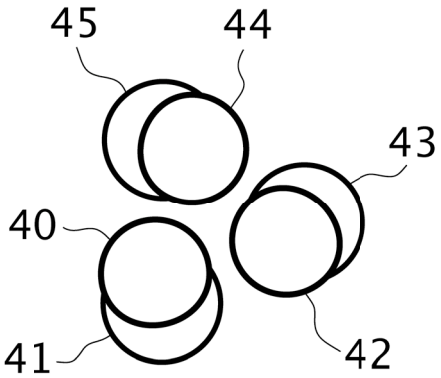

FIGURE 10

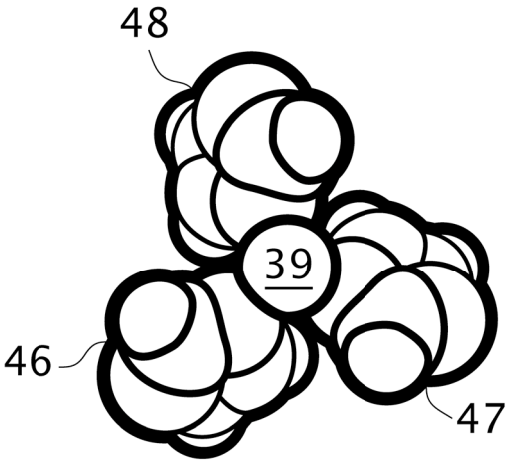

FIGURE 11

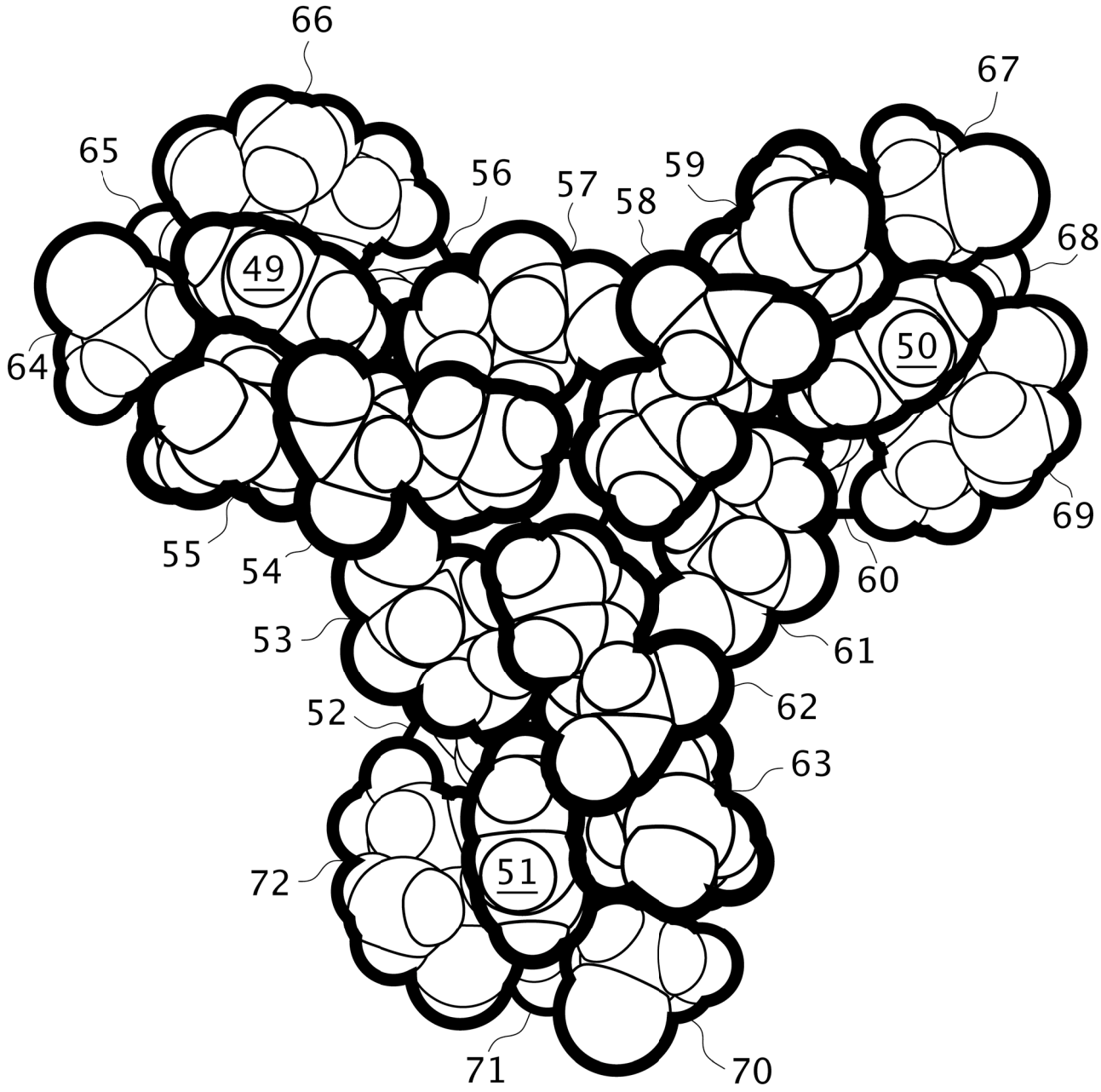

FIGURE 12

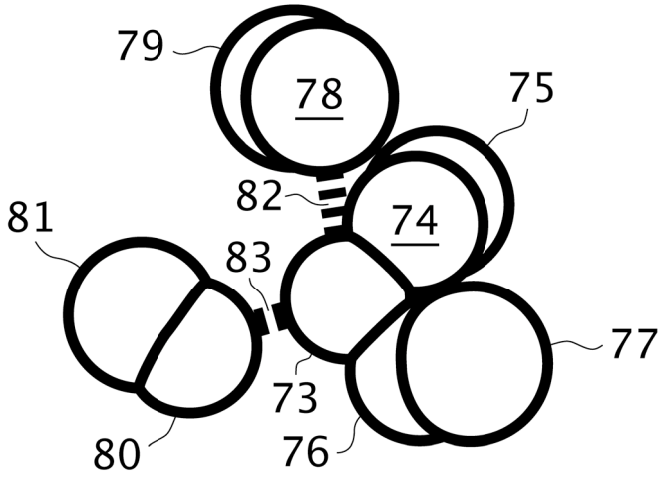

FIGURE 13

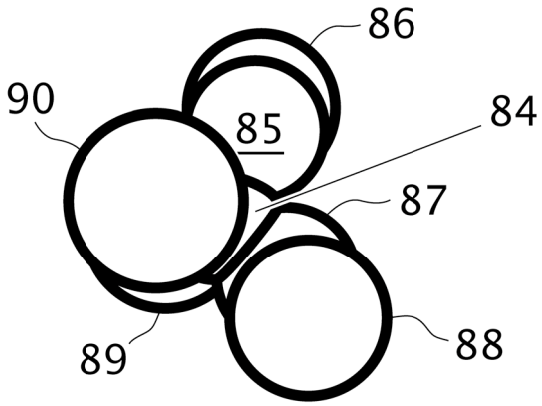

FIGURE 14

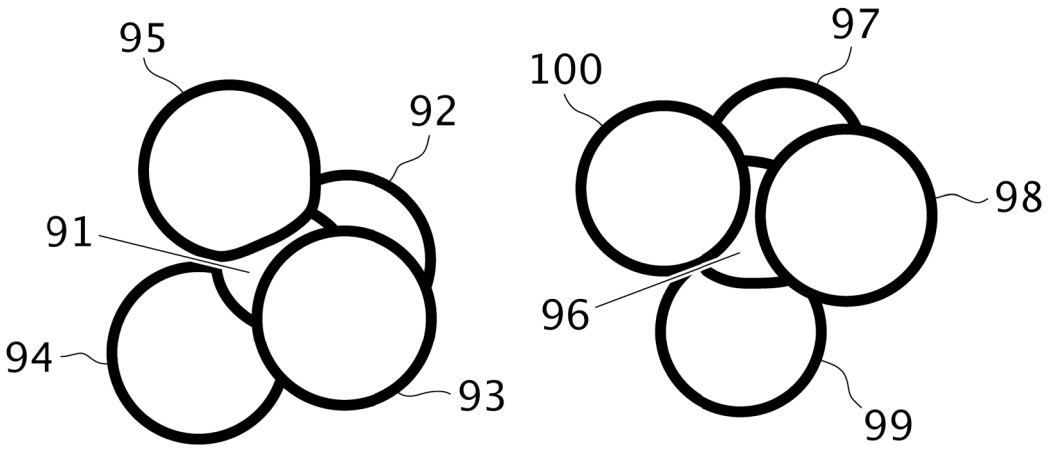

FIGURE 15

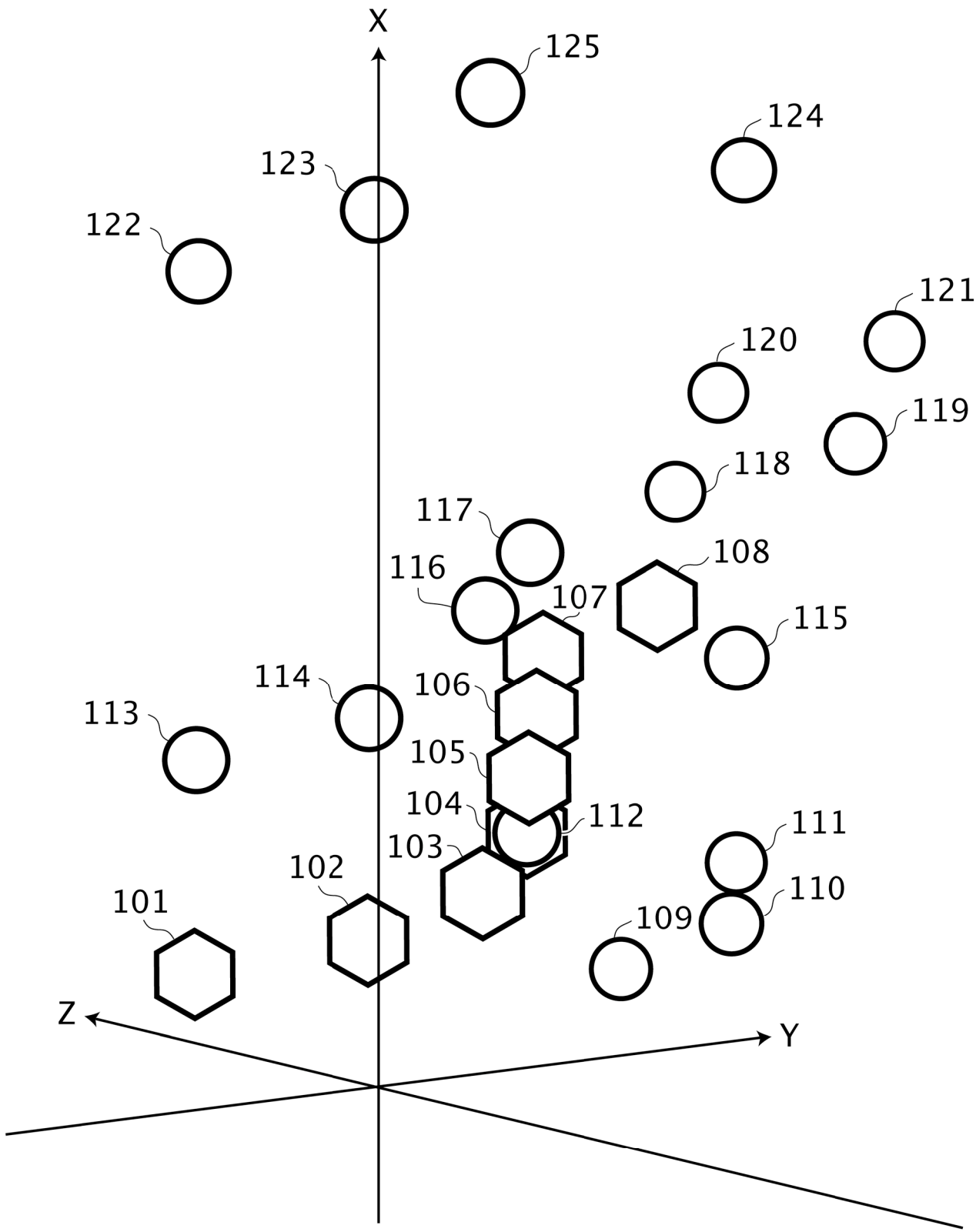
X
125
124
123
122
121
120
119
118
117
116
108
107
115
114
106
113
105
104
112
111
103
102
110
101
109
Z
Y

FIGURE 16

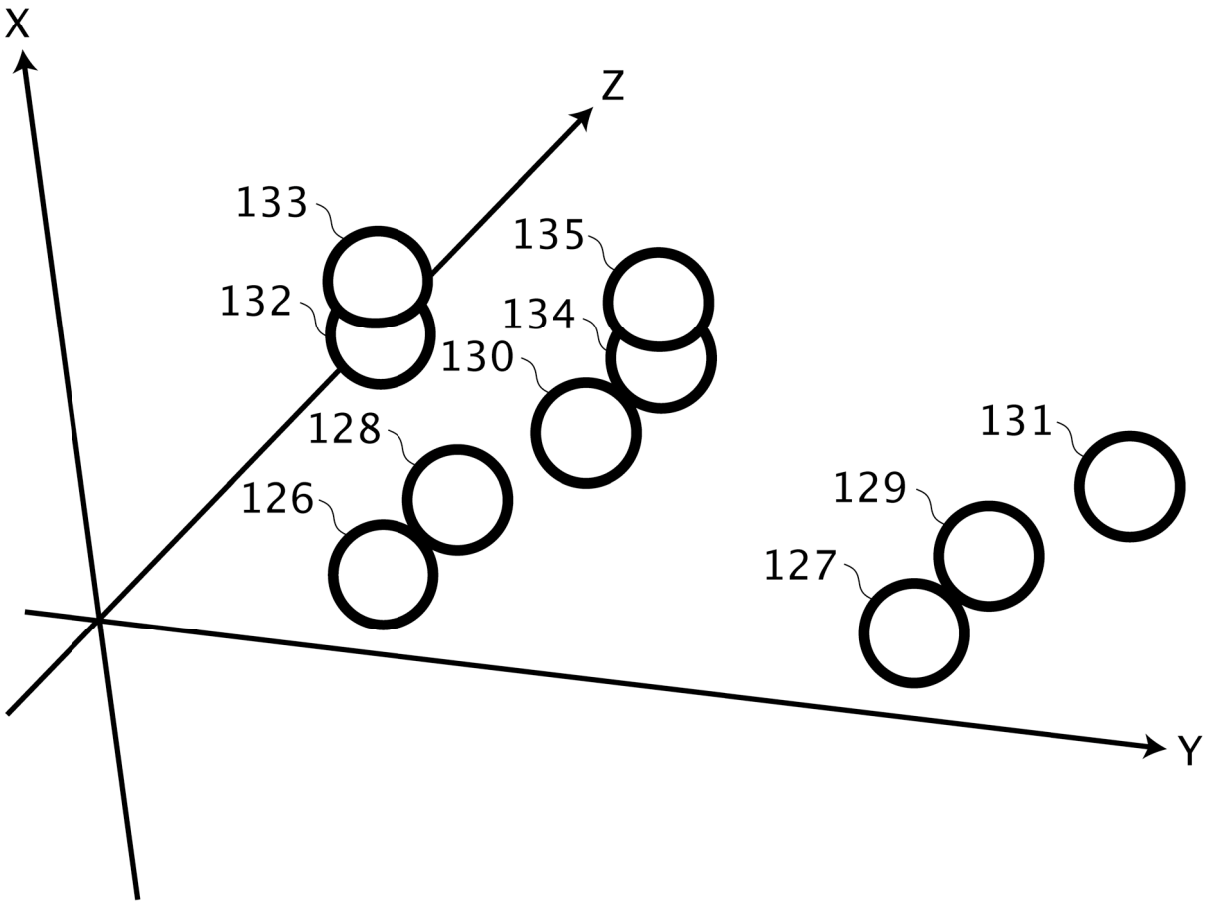

FIGURE 17

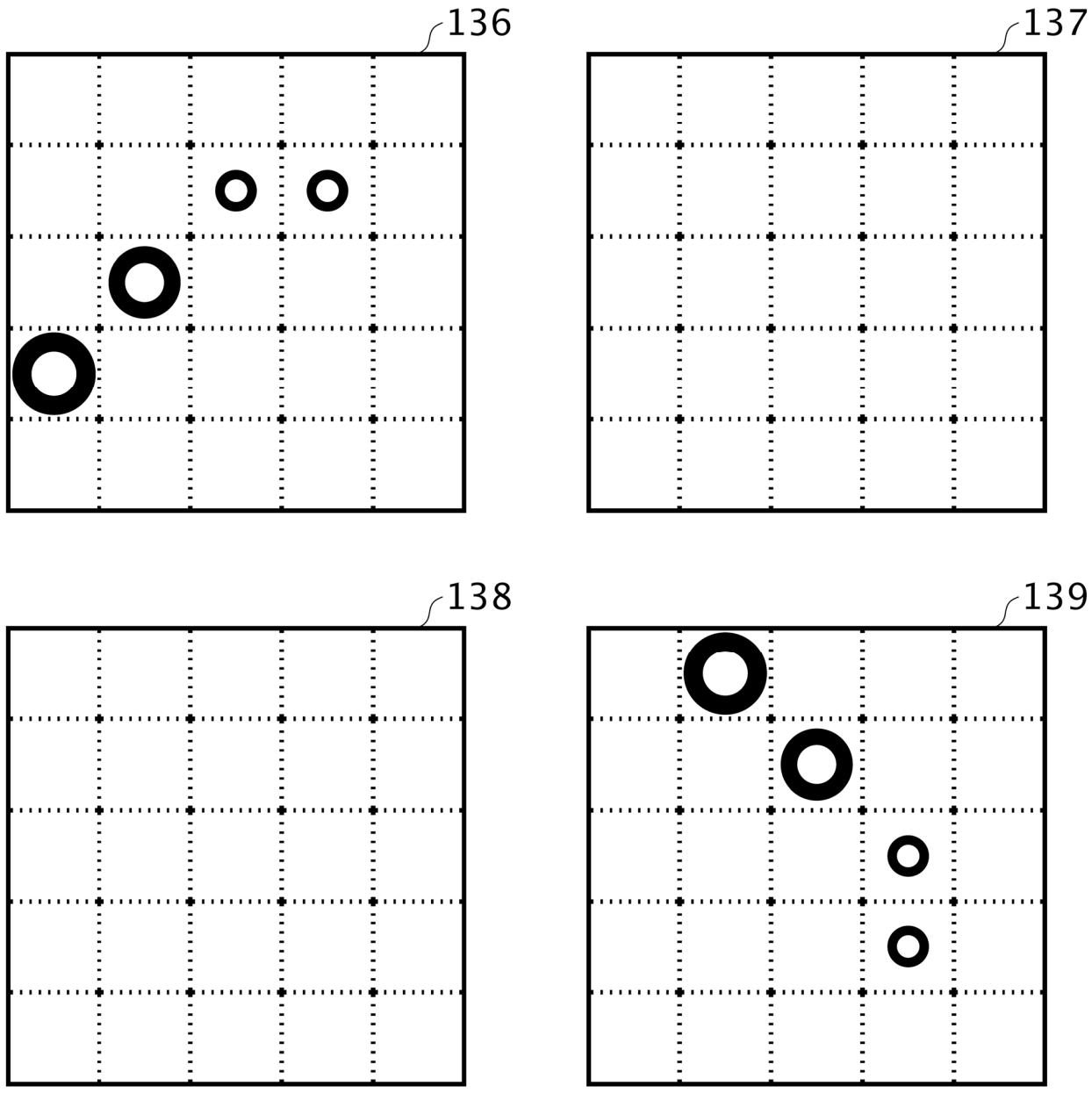

FIGURE 18

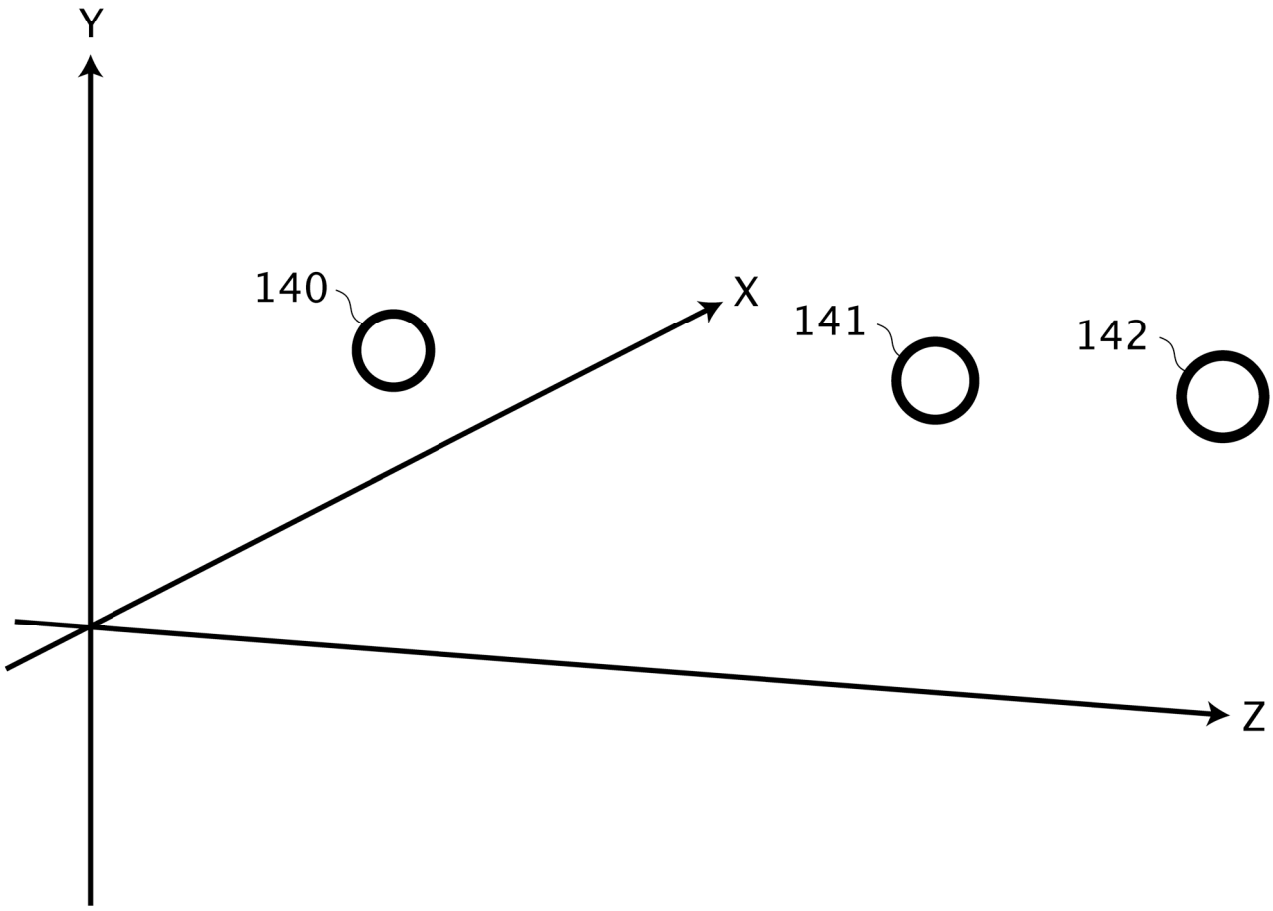

FIGURE 19