

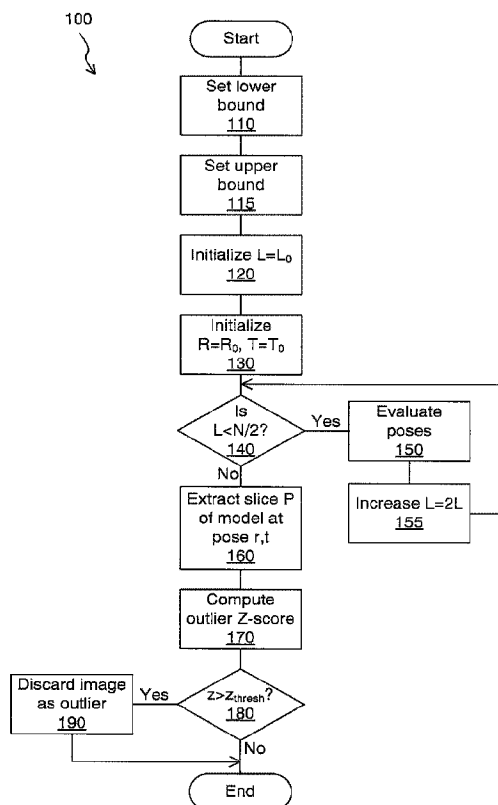


(86) Date de dépôt PCT/PCT Filing Date: 2017/05/16
(87) Date publication PCT/PCT Publication Date: 2017/11/23
(45) Date de délivrance/Issue Date: 2023/08/01
(85) Entrée phase nationale/National Entry: 2018/11/08
(86) N° demande PCT/PCT Application No.: CA 2017/050591
(87) N° publication PCT/PCT Publication No.: 2017/197514
(30) Priorité/Priority: 2016/05/16 (US62/336,831)

(51) Cl.Int./Int.Cl. *G06T 7/30* (2017.01)
(72) Inventeurs/Inventors:
PUNJANI, ALI, CA;
BRUBAKER, MARCUS ANTHONY, CA;
FLEET, DAVID JAMES, CA
(73) Propriétaire/Owner:
THE GOVERNING COUNCIL OF THE UNIVERSITY OF
TORONTO, CA
(74) Agent: Bhole IP Law

(54) Titre : PROCÉDES ET SYSTÈMES DESTINÉS À L'ALIGNEMENT D'IMAGE D'AU MOINS UNE IMAGE AVEC UN
MODELE

(54) Title: METHODS AND SYSTEMS FOR IMAGE ALIGNMENT OF AT LEAST ONE IMAGE TO A MODEL



(57) Abrégé/Abstract:

A system and a method for image alignment between at least two images to a three-dimensional model. The method including: determining a lower bound and an upper bound of an acceptable likelihood of mismatch between the at least two images; evaluating the likelihood of mismatch between the at least two images over a set of poses (r), shifts (t), or both poses (r) and shifts (t); and discarding those evaluations resulting beyond the lower bound and upper bound.

ABSTRACT

A system and a method for image alignment between at least two images to a three-dimensional model. The method including: determining a lower bound and an upper bound of an acceptable likelihood of mismatch between the at least two images; evaluating the likelihood of mismatch between the at least two images over a set of poses (r), shifts (t), or both poses (r) and shifts (t); and discarding those evaluations resulting beyond the lower bound and upper bound.

1 METHODS AND SYSTEMS FOR IMAGE ALIGNMENT OF AT LEAST ONE IMAGE TO A
2 MODEL

3 TECHNICAL FIELD

4 [0001] The following relates generally to image alignment and more particularly to aligning
5 images that share a characteristic using an iterative process.

6 SUMMARY

7 [0002] In one aspect, a method for image alignment of at least one two-dimensional or
8 three-dimensional image to a two-dimensional or three-dimensional model is provided, the
9 method executed on a processing unit, the method comprising: determining a lower bound and
10 an upper bound of an acceptable likelihood of mismatch between the at least one image and the
11 model; evaluating a value for the lower bound and the upper bound over a set of poses (r), shifts
12 (t), or both poses (r) and shifts (t); and discarding those evaluations where the lower bound is
13 greater than the upper bound.

14 [0003] In a particular case, the lower bound is determined with the images at a resolution
15 that is less than the maximum resolution for the images.

16 [0004] In a particular case, the determination of the lower bound comprises a combination of
17 a first component ($U(r, t)$) and a second component ($V(r, t)$).

18 [0005] In a particular case, the first component is the squared error of Fourier coefficients at
19 or below a selected radius in Fourier space, and the second component is the squared error of
20 Fourier coefficients above the selected radius.

21 [0006] In a particular case, the second component comprises a first subcomponent added to
22 a second subcomponent and subtracted by a third subcomponent, the first subcomponent is the
23 power of one of the images at high frequencies, the second subcomponent is the power of an
24 image of a slice of the three-dimensional model for one of the poses at high frequencies, the third
25 subcomponent is the correlation between the power of each of the images at high frequencies
26 and the power of the image of the slice of the three-dimensional model for one of the poses at
27 high frequencies.

28 [0007] In a particular case, the second subcomponent comprises: $V_1 - \sum_{\|l\| > L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2 -$
29 $4 \sqrt{\sum_{\|l\| > L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2}$, wherein V_1 is the power of one of the images at high frequencies, subscript
30 l denotes a wavevector, subscript L denotes the selected radius in the Fourier space, C is a

1 contrast transfer function (CTF) of the image-capturing apparatus, and Y is a vector representing
2 a projection of the three-dimensional model.

3 [0008] In a particular case, the second subcomponent is only recomputed if the CTF of the
4 image-capturing apparatus is different.

5 [0009] In a particular case, the determination of the upper bound comprises evaluating a
6 value for the likelihood of mismatch at a specific pose, specific shift, or both.

7 [0010] In a particular case, evaluating the value of the lower bound and the upper bound is
8 over a set of both poses and shifts, the method further comprising determining a shift and a pose
9 with favorable image alignment, by: selecting an initial value for a radius in the Fourier space;
10 discretizing the set of poses into a discrete grid of candidate poses and the set of shifts into a
11 discrete grid of candidate shifts; determining whether a selected accuracy of image alignment
12 has been obtained, when false: determining the lower bound for likelihood of mismatch for each
13 of the poses on the discrete grid of candidate poses and each of the shifts on the discrete grid of
14 candidate shifts; selecting the pose and shift where the lower bound has a minimum value;
15 determining the upper bound value, equal to the value of the likelihood of mismatch at the
16 selected pose and shift; for every pose in the discrete grid of candidate poses, determining
17 whether the minimum value of the lower bound over all candidate shifts is greater than the value
18 of the upper bound, and discarding the pose from the discrete grid in such case; for every shift in
19 the discrete grid of candidate shifts, determining whether the minimum value of the lower bound
20 over all candidate poses is greater than the value of the upper bound, and discarding the shift
21 from the discrete grid in such case; and increasing the radius in the Fourier space; and
22 otherwise, returning the pose and shift where the lower bound has the minimum value.

23 [0011] In a particular case, the selected accuracy of image alignment is obtained when the
24 value for the radius in Fourier space is equal to the Nyquist rate.

25 [0012] In a particular case, the selected accuracy of image alignment is obtained when a
26 selected number of iterations have been performed.

27 [0013] In a particular case, evaluating the value of the lower bound and the upper bound is
28 over a set of both poses and shifts, the method further comprising determining a shift and a pose
29 with favorable image alignment, by: selecting a value for a radius in Fourier space, the selected
30 value for the radius having an estimated low value relative to the possible values for the radius;
31 discretizing the set of poses into a discrete grid of candidate poses and the shifts into a discrete
32 grid of candidate shifts; determining whether a selected accuracy of image alignment has been

1 obtained, when false: determining the lower bound for likelihood of mismatch for each of the
 2 poses on the discrete grid of candidate poses and each of the shifts on the discrete grid of
 3 candidate shifts; selecting the pose and shift where the lower bound has a minimum value;
 4 determining the upper bound value, equal to the value of the likelihood of mismatch at the
 5 selected pose and shift; for every pose in the discrete grid of candidate poses, determining
 6 whether the minimum value of the lower bound over all candidate shifts is greater than the value
 7 of the upper bound, and discarding the pose from the discrete grid in such case; for every shift in
 8 the discrete grid of candidate shifts, determining whether the minimum value of the lower bound
 9 over all candidate poses is greater than the value of the upper bound, and discarding the shift
 10 from the discrete grid in such case; for every remaining pose in the discrete grid of candidate
 11 poses, replacing the pose with a plurality of subdivided grid points representing the candidate
 12 poses; for every remaining shift in the discrete grid of candidate shifts, replacing the shift with a
 13 plurality of subdivided grid points representing the candidate shifts; and increasing the radius in
 14 the Fourier space; and otherwise, returning the pose and shift where the lower bound has the
 15 minimum value.

16 [0014] In a particular case, the at least one image is two-dimensional and the model is
 17 three-dimensional, wherein replacing the pose with the plurality of subdivided grid points
 18 comprises replacing the pose with eight subdivided grid points, and wherein replacing the shift
 19 with the plurality of subdivided grid points comprises replacing the shift with four subdivided grid
 20 points.

21 [0015] In a particular case, the at least one image is two-dimensional and the model is
 22 two-dimensional, wherein replacing the pose with the plurality of subdivided grid points
 23 comprises replacing the pose with two subdivided grid points, and wherein replacing the shift
 24 with the plurality of subdivided grid points comprises replacing the shift with four subdivided grid
 25 points.

26 [0016] In a particular case, the at least one image is three-dimensional and the model is
 27 three-dimensional, wherein replacing the pose with the plurality of subdivided grid points
 28 comprises replacing the pose with eight subdivided grid points, and wherein replacing the shift
 29 with the plurality of subdivided grid points comprises replacing the shift with eight subdivided grid
 30 points.

31 [0017] In a particular case, the selected accuracy of image alignment is obtained when the
 32 value for the radius in Fourier space is equal to the Nyquist rate.

1 [0018] In a particular case, the selected accuracy of image alignment is obtained when a
2 selected number of iterations have been performed.

3 [0019] In a particular case, the method further comprises determining a shift (r) and a pose
4 (t) with favorable image alignment, by: selecting a value for a radius in Fourier space; discretizing
5 the set of poses into a discrete grid of candidate poses and the shifts into a discrete grid of
6 candidate shifts; determining a fixed fraction (f_{keep}) for the upper bound, the fixed fraction being
7 determined based on, at least, the fraction of poses and shifts that are discarded on typical
8 dataset of images; determining whether a selected accuracy of image alignment has been
9 obtained, when false: determining values for the first component for each of the poses and shifts
10 on the discrete grid of candidate poses and the discrete grid of shifts, using only portions of the
11 image below the selected value for the radius in Fourier space; determining a reference first
12 component (U^*) using the value for the fixed fraction; for every pose in the discrete grid of
13 candidate poses, determining whether the minimum value of the first component over all
14 candidate shifts is greater than the reference first component and discarding the pose from the
15 discrete grid in such case; for every shift in the discrete grid of candidate shifts, determining
16 whether the minimum value of the first component over all candidate poses is greater than the
17 reference first component and discarding the shift from the discrete grid in such case; for every
18 remaining pose in the discrete grid of candidate poses, replacing the pose with a plurality of
19 subdivided grid points representing the candidate poses; and for every remaining shift in the
20 discrete grid of candidate shifts, replacing the shift with a plurality of subdivided grid points
21 representing the candidate shifts; increasing the radius in the Fourier space; otherwise, returning
22 the pose and shift at the lower bound with minimum value.

23 [0020] In a particular case, determining the reference first component (U^*) comprises
24 determining the reference first component (U^*) such that: $\frac{|\{(r,t); U(r,t) \leq U^*\}|}{|\{(r,t); U(r,t) > U^*\}|} = f_{\text{keep}}$.

25 [0021] In a particular case, the fixed fraction (f_{keep}) is between 3% and 10%.

26 [0022] In a particular case, the fixed fraction (f_{keep}) is approximately 5%.

27 [0023] In a particular case, the at least one image is two-dimensional and the model is
28 three-dimensional, wherein replacing the pose with the plurality of subdivided grid points
29 comprises replacing the pose with eight subdivided grid points, and wherein replacing the shift
30 with the plurality of subdivided grid points comprises replacing the shift with four subdivided grid
31 points.

1 [0024] In a particular case, the at least one image is two-dimensional and the model is
 2 two-dimensional, wherein replacing the pose with the plurality of subdivided grid points
 3 comprises replacing the pose with two subdivided grid points, and wherein replacing the shift
 4 with the plurality of subdivided grid points comprises replacing the shift with four subdivided grid
 5 points.

6 [0025] In a particular case, the at least one image is three-dimensional and the model is
 7 three-dimensional, wherein replacing the pose with the plurality of subdivided grid points
 8 comprises replacing the pose with eight subdivided grid points, and wherein replacing the shift
 9 with the plurality of subdivided grid points comprises replacing the shift with eight subdivided grid
 10 points.

11 [0026] In a particular case, the selected accuracy of image alignment is obtained when the
 12 value for the radius in Fourier space is equal to the Nyquist rate.

13 [0027] In a particular case, the selected accuracy of image alignment is obtained when a
 14 selected number of iterations have been performed.

15 [0028] In another aspect, a system comprising a processor and a memory is provided, the
 16 memory having stored thereon computer instructions which when executed by the processor
 17 cause the processor to implement image alignment of at least one two-dimensional or
 18 three-dimensional image to a two-dimensional or three-dimensional model comprising:
 19 determining a lower bound and an upper bound of an acceptable likelihood of mismatch between
 20 the at least one image and the model; evaluating a value for the lower bound and the upper
 21 bound over a set of poses (r), shifts (t), or both poses (r) and shifts (t); and discarding those
 22 evaluations where the lower bound is greater than the upper bound.

23 [0029] In a particular case, the lower bound is determined with the images at a resolution
 24 that is less than the maximum resolution for the images.

25 [0030] In a particular case, the determination of the lower bound comprises a combination of
 26 a first component ($U(r, t)$) and a second component ($V(r, t)$).

27 [0031] In a particular case, the first component is the squared error of Fourier coefficients at
 28 or below a selected radius in Fourier space, and the second component is the squared error of
 29 Fourier coefficients above the selected radius.

30 [0032] In a particular case, the second component comprises a first subcomponent added to
 31 a second subcomponent and subtracted by a third subcomponent, the first subcomponent is the

1 power of one of the images at high frequencies, the second subcomponent is the power of an
 2 image of a slice of the three-dimensional model for one of the poses at high frequencies, the third
 3 subcomponent is the correlation between the power of each of the images at high frequencies
 4 and the power of the image of the slice of the three-dimensional model for one of the poses at
 5 high frequencies.

6 [0033] In a particular case, the second subcomponent comprises: $V_1 - \sum_{\|l\| > L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2 -$
 7 $4 \sqrt{\sum_{\|l\| > L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2}$, wherein V_1 is the power of one of the images at high frequencies, subscript
 8 l denotes a wavevector, subscript L denotes the selected radius in the Fourier space, C is a
 9 contrast transfer function (CTF) of the image-capturing apparatus, and Y is a vector representing
 10 a projection of the three-dimensional model.

11 [0034] In a particular case, the determination of the upper bound comprises evaluating a
 12 value for the likelihood of mismatch at a specific pose, specific shift, or both.

13 [0035] In a particular case, evaluating the value of the lower bound and the upper bound is
 14 over a set of both poses and shifts, the image alignment further comprising determining a shift
 15 and pose with favorable image alignment, by: selecting an initial value for a radius in the Fourier
 16 space; discretizing the set of poses into a discrete grid of candidate poses and the set of shifts
 17 into a discrete grid of candidate shifts; determining whether a selected accuracy of image
 18 alignment has been obtained, when false: determining the lower bound for likelihood of
 19 mismatch for each of the poses on the discrete grid of candidate poses and each of the shifts on
 20 the discrete grid of candidate shifts; selecting the pose and shift where the lower bound has a
 21 minimum value; determining the upper bound value, equal to the value of the likelihood of
 22 mismatch at the selected pose and shift; for every pose in the discrete grid of candidate poses,
 23 determining whether the minimum value of the lower bound over all candidate shifts is greater
 24 than the value of the upper bound, and discarding the pose from the discrete grid in such case;
 25 for every shift in the discrete grid of candidate shifts, determining whether the minimum value of
 26 the lower bound over all candidate poses is greater than the value of the upper bound, and
 27 discarding the shift from the discrete grid in such case; and increasing the radius in the Fourier
 28 space; and otherwise, returning the pose and shift where the lower bound has the minimum
 29 value.

30 [0036] In a particular case, the selected accuracy of image alignment is obtained when the
 31 value for the radius in Fourier space is equal to the Nyquist rate.

1 [0037] In a particular case, the selected accuracy of image alignment is obtained when a
2 selected number of iterations have been performed.

3 [0038] In a particular case, evaluating the value of the lower bound and the upper bound is
4 over a set of both poses and shifts, the image alignment further comprising: selecting a value for
5 a radius in Fourier space, the selected value for the radius having an estimated low value relative
6 to the possible values for the radius; discretizing the set of poses into a discrete grid of candidate
7 poses and the shifts into a discrete grid of candidate shifts; determining whether a selected
8 accuracy of image alignment has been obtained, when false: determining the lower bound for
9 likelihood of mismatch for each of the poses on the discrete grid of candidate poses and each of
10 the shifts on the discrete grid of candidate shifts; selecting the pose and shift where the lower
11 bound has a minimum value; determining the upper bound value, equal to the value of the
12 likelihood of mismatch at the selected pose and shift; for every pose in the discrete grid of
13 candidate poses, determining whether the minimum value of the lower bound over all candidate
14 shifts is greater than the value of the upper bound, and discarding the pose from the discrete grid
15 in such case; for every shift in the discrete grid of candidate shifts, determining whether the
16 minimum value of the lower bound over all candidate poses is greater than the value of the upper
17 bound, and discarding the shift from the discrete grid in such case; for every remaining pose in
18 the discrete grid of candidate poses, replacing the pose with a plurality of subdivided grid points
19 representing the candidate poses; for every remaining shift in the discrete grid of candidate
20 shifts, replacing the shift with a plurality of subdivided grid points representing the candidate
21 shifts; and increasing the radius in the Fourier space; and otherwise, returning the pose and shift
22 at the lower bound with minimum value.

23 [0039] In a particular case, the at least one image is two-dimensional and the model is
24 three-dimensional, wherein replacing the pose with the plurality of subdivided grid points
25 comprises replacing the pose with eight subdivided grid points, and wherein replacing the shift
26 with the plurality of subdivided grid points comprises replacing the shift with four subdivided grid
27 points.

28 [0040] In a particular case, the at least one image is two-dimensional and the model is
29 two-dimensional, wherein replacing the pose with the plurality of subdivided grid points
30 comprises replacing the pose with two subdivided grid points, and wherein replacing the shift
31 with the plurality of subdivided grid points comprises replacing the shift with four subdivided grid
32 points.

1 [0041] In a particular case, the at least one image is three-dimensional and the model is
 2 three-dimensional, wherein replacing the pose with the plurality of subdivided grid points
 3 comprises replacing the pose with eight subdivided grid points, and wherein replacing the shift
 4 with the plurality of subdivided grid points comprises replacing the shift with eight subdivided grid
 5 points.

6 [0042] In a particular case, the selected accuracy of image alignment is obtained when the
 7 value for the radius in Fourier space is equal to the Nyquist rate.

8 [0043] In a particular case, the selected accuracy of image alignment is obtained when a
 9 selected number of iterations have been performed.

10 [0044] In a particular case, evaluating the value of the lower bound and the upper bound is
 11 over a set of both poses and shifts, the image alignment further comprising: selecting a value for
 12 a radius in Fourier space; discretizing the set of poses into a discrete grid of candidate poses and
 13 the shifts into a discrete grid of candidate shifts; determining a fixed fraction (f_{keep}) for the upper
 14 bound, the fixed fraction being determined based on, at least, the fraction of poses and shifts that
 15 are discarded on typical dataset of images; determining whether a selected accuracy of image
 16 alignment has been obtained, when false: determining values for the first component for each of
 17 the poses and shifts on the discrete grid of candidate poses and the discrete grid of shifts, using
 18 only portions of the image below the selected value for the radius in Fourier space; determining a
 19 reference first component (U^*) using the value for the fixed fraction; for every pose in the
 20 discrete grid of candidate poses, determining whether the minimum value of the first component
 21 over all candidate shifts is greater than the reference first component and discarding the pose
 22 from the discrete grid in such case; for every shift in the discrete grid of candidate shifts,
 23 determining whether the minimum value of the first component over all candidate poses is
 24 greater than the reference first component and discarding the shift from the discrete grid in such
 25 case; for every remaining pose in the discrete grid of candidate poses, replacing the pose with a
 26 plurality of subdivided grid points representing the candidate poses; and for every remaining shift
 27 in the discrete grid of candidate shifts, replacing the shift with a plurality of subdivided grid points
 28 representing the candidate shifts; increasing the radius in the Fourier space; otherwise, returning
 29 the pose and shift at the lower bound with minimum value.

30 [0045] In a particular case, determining the reference first component (U^*) comprises
 31 determining the reference first component (U^*) such that: $\frac{| \{(r,t); U(r,t) \leq U^* \} |}{| \{(r,t); U(r,t) > U^* \} |} = f_{\text{keep}}$.

1 [0046] In a particular case, the at least one image is two-dimensional and the model is
 2 three-dimensional, wherein replacing the pose with the plurality of subdivided grid points
 3 comprises replacing the pose with eight subdivided grid points, and wherein replacing the shift
 4 with the plurality of subdivided grid points comprises replacing the shift with four subdivided grid
 5 points.

6 [0047] In a particular case, the at least one image is two-dimensional and the model is
 7 two-dimensional, wherein replacing the pose with the plurality of subdivided grid points
 8 comprises replacing the pose with two subdivided grid points, and wherein replacing the shift
 9 with the plurality of subdivided grid points comprises replacing the shift with four subdivided grid
 10 points.

11 [0048] In a particular case, the at least one image is three-dimensional and the model is
 12 three-dimensional, wherein replacing the pose with the plurality of subdivided grid points
 13 comprises replacing the pose with eight subdivided grid points, and wherein replacing the shift
 14 with the plurality of subdivided grid points comprises replacing the shift with eight subdivided grid
 15 points.

16 [0049] In a particular case, the selected accuracy of image alignment is obtained when the
 17 value for the radius in Fourier space is equal to the Nyquist rate.

18 [0050] In a particular case, the selected accuracy of image alignment is obtained when a
 19 selected number of iterations have been performed.

20 DESCRIPTION OF THE DRAWINGS

21 [0051] The features of the invention will become more apparent in the following detailed
 22 description in which reference is made to the appended drawings wherein:

23 [0052] Fig. 1 is a system for image alignment in accordance with an embodiment;

24 [0053] Fig. 2 is a flow chart of the general method for image alignment used by the system of
 25 Fig. 1;

26 [0054] Figs. 3A to 3C show three slices from the x, y, and z directions of a typical model of a
 27 protein in Cryo-EM (TRPA1);

28 [0055] Fig. 4 shows a falloff of model power as a function of frequency;

29 [0056] Fig. 5A shows an image taken with a Cryo-EM of a protein;

30 [0057] Figs. 5B to 5D show the image of 5A after application of various filters;

[0058] Fig. 6 shows the values of a bound over poses r , taking the minimum over all shifts to arrive at the values for each r ;

[0059] Fig. 7 shows a zoomed in section of the graph of Fig. 6 near the true best pose;

[0060] Fig. 8 shows the same bounds as Figs. 6 and 7, but over shifts rather than poses;

[0061] Fig. 9 shows the process of evaluating poses of Fig. 2 in accordance with a first implementation;

[0062] Fig. 10 shows the process of evaluating poses of Fig. 2 in accordance with a second implementation that uses first level approximations;

[0063] Fig. 11 shows the process of evaluating poses of Fig. 2 in accordance with a third implementation that uses second level approximations;

[0064] Figs. 12A to 12C show three slices from the x, y, and z directions of a protein reconstruction result in Cryo-EM (TRPV1);

[0065] Fig. 13 is an FSC resolution estimate of the protein of Figs. 12A to 12C;

[0066] Fig. 14 is a rendered image of the reconstruction of the determined structure(s) of Figs. 12A to 12C;

[0067] Figs. 15A to 15C show three slices from the x, y, and z directions of a reconstructed ribosome protein in Cryo-EM; and

[0068] Fig. 16 is an FSC resolution estimate of the ribosome of Figs. 15A to 15C.

DETAILED DESCRIPTION

[0069] For simplicity and clarity of illustration, where considered appropriate, reference numerals may be repeated among the Figures to indicate corresponding or analogous elements. In addition, numerous specific details are set forth in order to provide a thorough understanding of the embodiments described herein. However, it will be understood by those of ordinary skill in the art that the embodiments described herein may be practiced without these specific details. In other instances, well-known methods, procedures and components have not been described in detail so as not to obscure the embodiments described herein. Also, the description is not to be considered as limiting the scope of the embodiments described herein.

[0070] Various terms used throughout the present description may be read and understood as follows, unless the context indicates otherwise: “or” as used throughout is inclusive, as though written “and/or”; singular articles and pronouns as used throughout include their plural forms, and

1 vice versa; similarly, gendered pronouns include their counterpart pronouns so that pronouns
2 should not be understood as limiting anything described herein to use, implementation,
3 performance, etc. by a single gender; "exemplary" should be understood as "illustrative" or
4 "exemplifying" and not necessarily as "preferred" over other embodiments. Further definitions for
5 terms may be set out herein; these may apply to prior and subsequent instances of those terms,
6 as will be understood from a reading of the present description.

7 [0071] Any module, unit, component, server, computer, terminal, engine or device
8 exemplified herein that executes instructions may include or otherwise have access to computer
9 readable media such as storage media, computer storage media, or data storage devices
10 (removable and/or non-removable) such as, for example, magnetic disks, optical disks, or tape.
11 Computer storage media may include volatile and non-volatile, removable and non-removable
12 media implemented in any method or technology for storage of information, such as computer
13 readable instructions, data structures, program modules, or other data. Examples of computer
14 storage media include RAM, ROM, EEPROM, flash memory or other memory technology,
15 CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic
16 tape, magnetic disk storage or other magnetic storage devices, or any other medium which can
17 be used to store the desired information and which can be accessed by an application, module,
18 or both. Any such computer storage media may be part of the device or accessible or
19 connectable thereto. Further, unless the context clearly indicates otherwise, any processor or
20 controller set out herein may be implemented as a singular processor or as a plurality of
21 processors. The plurality of processors may be arrayed or distributed, and any processing
22 function referred to herein may be carried out by one or by a plurality of processors, even though
23 a single processor may be exemplified. Any method, application or module herein described may
24 be implemented using computer readable/executable instructions that may be stored or
25 otherwise held by such computer readable media and executed by the one or more processors.

26 [0072] Image alignment challenges occur in several fields of use. As an example, in Single
27 Particle Electron Cryomicroscopy (Cryo-EM), scientists aim to discover the structure and function
28 of important biological macromolecules (proteins, viruses, complexes, etc) by computing 3D
29 atomic structures from a large set of noisy 2D electron micrograph images of the target molecule.
30 Cryo-EM has the potential to unveil the molecular and chemical nature of fundamental biology
31 through the discovery of atomic structures of previously unknown biological structures, many of
32 which have proven difficult or impossible to study by conventional structural biology techniques.
33 As such, Cryo-EM is receiving considerable attention in the life science and structural biology

1 fields. Unfortunately, progress in employing Cryo-EM to study biology widely is hindered by the
 2 difficulty of the associated computational task. Today, biologists consider computational
 3 expense, in terms of both cost for compute infrastructure and more importantly time wasted
 4 waiting for computation, to be a key bottleneck in the Cryo-EM pipeline. Current state-of-the-art
 5 methods require weeks or even months of compute time on large clusters with several
 6 thousands of cores. In contrast, cryo-imaging of a specimen can currently be done within 48
 7 hours.

8 [0073] The power of cryo-EM to resolve the structures of complex proteins depends entirely
 9 on the power of techniques underlying the reconstruction process. Once a good quality signal
 10 can be recorded in the microscope, the rest of the burden to actually find a structure comes down
 11 to the techniques for processing resulting images. This is especially true as cryo-EM is
 12 increasingly applied to more and more difficult-to-purify proteins and heterogeneous samples
 13 where many of the collected microscope images do not actually contain the target particle. It is
 14 common to see collected datasets with as little as 10 or 20 percent true particles. This high
 15 outlier rate is further exacerbated by auto-picking procedures which aim to automate the manual
 16 process of finding particles in microscope images. For these reasons, it is clear that the
 17 technique for separating true particle images from false ones is critical to reliably discovering
 18 new structures, and also in making the process scalable and routine.

19 [0074] Reconstruction of a structure in Cryo-EM is generally considered to be expensive. In
 20 the Cryo-EM imaging process, each observed image represents a view of the target 3D structure
 21 that it is desired to recover. Crucially, each image comes from an unknown viewing direction. In
 22 modern approaches, this pose ambiguity is dealt with by formulating the reconstruction problem
 23 as an optimization problem over the space of 3D structures. The objective is to maximize the
 24 likelihood (or similar measure) of the structure, given the observed images. The pose variables
 25 are treated as nuisance (i.e., latent) variables. As such, computing the optimization objective
 26 necessitates a search problem, known as alignment. In this search problem, which must be
 27 carried out for each observed image, the latent pose variables are estimated by searching for the
 28 3D pose and 2D in-plane shift of the 3D structure that best explains the observed image. In some
 29 approaches, this search is done implicitly by marginalizing over the pose variable, but this still
 30 requires finding the (multitude of) poses for which a particular image aligns well.

31 [0075] $E(r, t)$ is used to denote the alignment error of an image to a 3D structure that has
 32 been rotated to the pose r (3D) and translated in-plane by shift t (2D). Alignment, or the
 33 minimization of $E(r, t)$, is expensive in current methods because it is solved using exhaustive

1 search over all poses and shifts. Persons of skill in the art are aware that misalignment of images
2 can lead to incorrect 3D structures, and so are keen to ensure that no poses are missed during
3 alignment. When attempting to align images so that atomic resolution structure can be
4 reconstructed, the five dimensional search space must be finely discretized, increasing
5 computational burden further still. Some methods apply ad-hoc measures to cut down on
6 computational cost; these are usually referred to as local searches. Current local search
7 methods come in many flavors, all being ad-hoc and giving no guarantee of finding the correct
8 alignment. Some are based on limiting searches to the region of a previously determined
9 orientation. Others use the concept of coarse-to-fine search, where an alignment is first
10 computed over a coarse grid of possible poses, and subsequently subsampled on finer grids.
11 Local searches are used in practice but only cautiously, as they can introduce alignment
12 inaccuracies and have no guarantees. They can also have tuning parameters that can be difficult
13 to set, as there is no mathematical backing for the ad-hoc procedures.

14 [0076] Disclosed herein are methods and systems for image alignment. The methods and
15 systems described herein are generally applicable to any set of images sharing a characteristic,
16 a specific example of which is Cryo-EM images, for which alignment is desired between various
17 2D images representing specific poses of a 3D specimen. Other examples of 2D/3D image
18 alignment have imaging examples deriving from microscopes, telescopes, NMR, CT scanners,
19 and MRI. Further examples include those for which 2D/2D and 3D/3D image alignment,
20 symmetry detection or sub-tomography is suitable.

21 [0077] The method comprises defining an upper bound and a lower bound delimiting an
22 acceptable likelihood of mismatch between two given images. In the specific example of
23 Cryo-EM, images are compared to each of a set of expected images of a protein structure at a
24 set of poses and shifts spaced apart at a first coarseness to determine a likelihood of a mismatch
25 between the image and the expected image. The poses and shifts at which the expected images
26 have likelihoods that exceed an upper bound are discarded. The process is repeated for a
27 second set of poses and shifts around the first set of poses and shifts corresponding to
28 un-discarded ones of the expected images, the second set of poses and shifts having a second
29 coarseness that is more fine than the first coarseness. The process is further repeated in further
30 iterations, with sets of poses and shifts becoming successively finer, while determining
31 likelihoods of mismatch at each iteration and discarding successively more poses and shifts, until
32 a suitable termination criterion is reached. The best of the remaining poses and shifts is
33 considered the correct alignment of the image to the reference.

1 [0078] Fig. 1 shows various physical components of a system 20 for image alignment. As will
2 be appreciated, while the system 20 is illustrated as being a single physical computer, it can
3 alternatively be two or more computers acting cooperatively to provide the functionality
4 described. As shown, the system 20 has a number of physical and logical components, including
5 a central processing unit ("CPU") 60, random access memory ("RAM") 64, an input/output ("I/O")
6 interface 68, a network interface 72, non-volatile storage 76, and a local bus 80 enabling the
7 CPU 60 to communicate with the other components. The CPU 60 executes an operating system
8 and an application for performing image alignment. The functionality of an application of the
9 image alignment method is described below in greater detail. The RAM 64 provides relatively
10 responsive volatile storage to the CPU 60. The I/O interface 68 enables an administrator to
11 interact with the system 20 via a keyboard, a mouse, a speaker, and a display. The network
12 interface 72 permits wired or wireless communication with other systems. The non-volatile
13 storage 76 stores computer readable instructions for implementing the operating system and the
14 application for performing image alignment, as well as any data used by the application. During
15 operation of the system 20, the computer readable instructions for the operating system and the
16 application, and the data may be retrieved from the non-volatile storage 76 and placed in the
17 RAM 64 to facilitate execution. An imaging system 74 may further be linked to the system to
18 obtain images for alignment. The imaging system comprise electron microscopes,
19 microscopes, telescopes, NMR, CT scanners, and MRI or other suitable device.

20 [0079] The system 20 executes a general optimization technique known as branch and
21 bound ("BnB") optimization. BnB optimization provides global optimality guarantees and, when
22 executed well, can lead to significant reductions in processing time. BnB operates generally as
23 follows. When searching for the minimum of some objective function, the search space is
24 subdivided (branch) and a method is determined for inexpensively discarding (bound) regions of
25 the search space that cannot possibly contain the minimum. The bound step of BnB requires
26 both an upper and lower bound on the objective function, and typically finding a method for
27 computing a useful lower bound is the most difficult. The bound, given a search region, is
28 designed to determine whether a specified region could possibly contain the minimum, and if not
29 then the region as a whole can be discarded. A powerful and useful lower bound has two
30 qualities: inexpensive computational requirements, since the evaluation is conducted many
31 times, and that the evaluation results in many discarded regions, meaning that the bound is tight
32 and actually provides useful information.

1 [0080] In the present system and method, a set of bounds for image alignment indicate, in a
 2 principled way, how poorly or well a given image can align to a given 3D map over a range of
 3 poses and shifts, without having to actually evaluate the exact fit error at those poses and shifts.
 4 This allows entire regions of pose and shift search space to be discarded using only inexpensive
 5 calculations, and allows the best pose to be found efficiently.

6 [0081] The problem of recovering a 3D cryo-EM structure of a molecule from a set of 2D
 7 images taken in a microscope is difficult for several reasons, key amongst them the presence of
 8 unobserved latent variables like pose and translational shift of the molecule in each image, a
 9 problem referred to as the latent variable estimation problem. A person of skill will appreciate the
 10 details of various techniques that can solve a latent variable problem, most of which comprise an
 11 iterative procedure whereby the current guess of the 3D structure is used to estimate the latent
 12 variables (pose, shift), and then these estimates are used to construct a new 'refined' guess of
 13 the 3D structure.

14 [0082] A method 100 for image alignment provided herein used by the system 20 in a first
 15 implementation, being aligning a 2D image from a Cryo-EM imaging system to a 3D protein
 16 structure, will now be described with reference to Fig. 2.

17 [0083] The general problem solved is computing and minimizing of the negative log
 18 probability of an image X (2D) given a model M (3D). This is a function of the latent variables
 19 describing the pose r and shift t . The likelihood is:

$$20 \quad p(X|M, r, t) = \frac{1}{Z} \exp \left(\sum_l \frac{-1}{2\sigma_l^2} |C_l Y_l(r) - S_l(t) X_l|^2 \right), \text{ where } Y_l(r) = \phi_l(r) M$$

21 [0084] Here, both image X and model M are represented in Fourier space. $Y(r)$ denotes
 22 the projection of model M along the pose given by r . Poses can be parameterized in any
 23 suitable fashion but the axis-angle formulation is applied in the current embodiment. The
 24 subscript l denotes a two-component index of a particular Fourier coefficient, also known as a
 25 wavevector. For instance $l = (0,0)$ would be the DC component, $l = (1,2)$ would be the Fourier
 26 component for wave with 1 cycle in the x -dimension and 2 cycles in the y -dimension. The
 27 sum over l is shorthand for summing over all wavevectors in 2D. C denotes the Contrast
 28 Transfer Function (CTF) of the microscope (or other image capturing apparatus), and $\phi_l(r)$ is a
 29 linear projection operator which in Fourier space is the slice operator with pose r , for wavevector
 30 l . S denotes the 2D phase shift corresponding to a shift of t (2D) pixels. Z can be ignored
 31 because it does not depend on r, t . The noise model σ_l represents the level of Gaussian noise
 32 expected in the images, with a possibly different variance for each Fourier coefficient (which

1 allows for white or colored noise models). For clarity in the following, σ_l is set to $\sigma_l = \sigma = 1$ but
 2 the general case would be apparent to a person of skill in the art. Next, the negative log is taken
 3 to arrive at our objective function which ends up being the squared error:

$$4 \quad E(r, t) = \sum_l \frac{1}{2} |C_l Y_l(r) - S_l(t) X_l|^2$$

5 [0085] The goal is to minimize $E(r, t)$ which depends on the image X and model M .

6 [0086] Current approaches to Cryo-EM alignment use exhaustive scan to search over poses
 7 and shifts. In exhaustive scan, a discrete set of poses $r \in \mathbf{R}$ and shifts $t \in \mathbf{T}$ are used and the
 8 minimum error pose on this discrete grid is found by directly evaluating $E(r, t)$ at each pose and
 9 shift and selecting the best one. In using exhaustive scan, an assumption is made that the
 10 sampling used in $\mathbf{R}$ and $\mathbf{T}$ is fine enough to capture the variations in $E(r, t)$, so that the best
 11 looking pose within the discrete grid will actually be close to the minimum of $E(r, t)$.

12 [0087] Unlike local searches, which can also be used to provide a modest speedup, the BnB
 13 technique described here is guaranteed to find the best pose in the sets $\mathbf{R}, \mathbf{T}$ without requiring
 14 an exhaustive scan.

15 [0088] The method 100 commences with the setting of a lower bound (110). The core
 16 difficulty in employing a BnB approach is to derive bounds that are cheap to evaluate but
 17 informative about the objective function $E(r, t)$. This usually requires insight into the
 18 characteristics of E .

19 [0089] The setting of the lower bound takes as an assumption that if a 2D image aligns
 20 poorly to a 3D structure at low resolution, it will probably not align well at high resolution. This
 21 means that if the likelihood of an image is evaluated using only its low resolution Fourier
 22 coefficients, this is likely to provide knowledge about which regions of pose and shift space are
 23 worth pursuing at high resolution.

24 [0090] The negative log-likelihood is a sum-of-squared-errors. This means that each Fourier
 25 coefficient is equivalent, and each contributes independently with equal weight to the total
 26 squared error E . The contribution of each Fourier coefficient is related to its power. If a Fourier
 27 coefficient in the model with wavevector l has no power, that coefficient will only contribute a
 28 term equal to $\frac{1}{2} |X_l|^2$ to E , and that term does not depend on the pose (r, t) and thus does not
 29 need to be considered during search. The bound described herein shows and exploits the fact
 30 that a Fourier coefficient in the model that has non-zero but small power also gives a small and
 31 limited possible pose-dependent contribution to E .

[0091] Therefore, if particular Fourier coefficients of the model at higher resolutions have limited power, there is a limit to how much they can impact the squared error E . If the low frequency coefficients already have a certain amount of error at a particular pose and shift, the high frequency coefficients cannot make this much better or worse. This sets a basis for constructing a bound on E that uses inexpensive evaluations of the squared error at low resolutions to bound true values of E , enabling the elimination of search regions after only evaluating them inexpensively at low resolution.

[0092] Figs. 3A to 3C show three slices from the x, y, and z directions of a typical model of a protein in Cryo-EM (TRPA1). Fig. 4 shows the falloff of model power as a function of frequency and shows that low frequencies have much more power than high frequencies.

[0093] It is desired to derive a lower bound that is always less than E . To start, E is split into two parts:

$$E(r, t) = \sum_{\|t\| \leq L} \frac{1}{2} |C_l Y_l(r) - S_l(t) X_l|^2 + \sum_{\|t\| > L} \frac{1}{2} |C_l Y_l(r) - S_l(t) X_l|^2 \quad (1)$$

$$E(r, t) = U(r, t) + V(r, t) \quad (2)$$

[0094] Thus, U is the squared error of Fourier coefficients at or below some radius in Fourier space L , and V is the squared error of coefficients above that radius.

[0095] In order to bound E , U is directly computed, which is cheap, but V is bounded from below. To do this, V is split into parts.

$$V(r, t) = \sum_{\|t\| > L} \frac{1}{2} |C_l Y_l(r) - S_l(t) X_l|^2 \quad (3)$$

$$= \sum_{\|t\| > L} \frac{1}{2} |X_l|^2 + \sum_{\|t\| > L} \frac{1}{2} C_l^2 |Y_l(r)|^2 - \sum_{\|t\| > L} \Re(C_l Y_l(r)^* S_l(t) X_l) \quad (4)$$

$$= V_1 + V_2 - V_3 \quad (5)$$

[0096] Here, the fact that $|S_l(t)| = 1$ is used, and also assumed that the CTF is real-valued. The first term, V_1 , is the power of the image at high frequencies, and does not depend on r, t . The second term is the power of a slice from pose r of the model at high frequencies. This does not depend on t . The third term is the correlation between the shifted image X and the slice of the model.

[0097] First V_3 is examined; an upper bound on V_3 contributes to a lower bound on V . The Cryo-EM image formation model is employed to start. This model says that the observed image X is a sum of a CTF-corrupted true signal plus independent, identically distributed noise. This

1 means X_l can be written as $X_l = C_l \bar{X}_l + \eta_l$ where all η_l 's are independent, identically
 2 distributed complex normal random variables with $\eta_l \sim \mathcal{CN}\left(0, \frac{\sigma_l^2}{2}\right)$ and it is assumed that $\sigma_l = 1$.

3 [0098] This results in

$$4 \quad V_3 = \sum_{\|l\| > L} C_l^2 \Re(Y_l(r)^* S_l(t) \bar{X}_l) + \sum_{\|l\| > L} \Re(C_l Y_l(r)^* S_l(t) \eta_l) \quad (6)$$

$$5 \quad = \sum_{\|l\| > L} C_l^2 \Re(Y_l(r)^* S_l(t) \bar{X}_l) + H \quad (7)$$

$$6 \quad \leq \sum_{\|l\| > L} C_l^2 |Y_l(r)| |\bar{X}_l| + H \quad (8)$$

7 [0099] Here, H is the random variable corresponding to the correlation between the model
 8 slice $Y(r)$ and the random noise η . H indicates how much to reasonably expect the noise to
 9 affect E . H can be written as:

$$10 \quad H = \sum_{\|l\| > L} \Re(C_l Y_l(r)^* S_l(t) \eta_l) \quad (9)$$

$$11 \quad = \sum_{\|l\| > L} C_l \Re(Y_l(r)^* \eta_l) \quad (10)$$

$$12 \quad = \sum_{\|l\| > L} C_l \Re\left(Y_l(r)^* \mathcal{CN}\left(0, \frac{\sigma^2}{2}\right)\right) \quad (11)$$

$$13 \quad = \sum_{\|l\| > L} C_l \Re\left(\mathcal{CN}\left(0, \frac{\sigma^2}{2} |Y_l(r)|^2\right)\right) \quad (12)$$

$$14 \quad = \sum_{\|l\| > L} \mathcal{N}\left(0, \frac{\sigma^2}{2} C_l^2 |Y_l(r)|^2\right) \quad (13)$$

$$15 \quad = \mathcal{N}\left(0, \sum_{\|l\| > L} \frac{\sigma^2}{2} C_l^2 |Y_l(r)|^2\right) \quad (14)$$

16 [0100] In the first line above, the noise variables η_l are uniform over phase, and so are
 17 invariant to phase shifting by $S_l(t)$. The final line means that H , the contribution to V from noise
 18 in the image, is normally distributed with the variance as given above. The standard deviation of
 19 H can be written as

$$20 \quad \sigma_H = \sqrt{\sum_{\|l\| > L} \frac{\sigma_l^2}{2} C_l^2 |Y_l(r)|^2}$$

21 [0101] Returning to V_2 and V_3 :

$$22 \quad V_2 - V_3 = \sum_{\|l\| > L} \frac{1}{2} C_l^2 |Y_l(r)|^2 - \sum_{\|l\| > L} \Re(C_l Y_l(r)^* S_l(t) X_l) \quad (15)$$

$$23 \quad V_2 - V_3 \geq \sum_{\|l\| > L} \frac{1}{2} C_l^2 |Y_l(r)|^2 - \sum_{\|l\| > L} C_l^2 |Y_l(r)| |\bar{X}_l| - H \quad (16)$$

$$= \sum_{\|l\| > L} \frac{1}{2} (C_l^2 |Y_l(r)|^2 - 2C_l^2 |Y_l(r)| |\tilde{X}_l|) - H \quad (17)$$

$$= \frac{1}{2} (\mathbf{Y}(r)^T \mathbf{D} \mathbf{Y}(r) - 2\mathbf{Y}(r)^T \mathbf{D} \tilde{\mathbf{X}}) - H \quad (18)$$

[0102] Equation 18 is written in vector form. $\mathbf{Y}(r)$ is a vector where each element is the absolute value of a Fourier coefficient $|Y_l(r)|$. Similarly $\tilde{\mathbf{X}}$ is a vector of $|\tilde{X}_l|$. $\mathbf{D}$ is a diagonal matrix of the CTF squared. Equation 18 is a quadratic form in $\mathbf{Y}(r)$. Turning again to the image formation model it is assumed that in the image, the true signal $\tilde{X}$ is actually a slice of the model in Fourier space. That means that $\tilde{X}$ is a slice of M from the true pose r^* ; $\tilde{X}_l = Y_l(r^*)$. Since Equation 18 is a quadratic form in $\mathbf{Y}(r)$ with a positive semi-definite Hessian $\mathbf{D}$, it can be said that the minimum of this expression with respect to $\mathbf{Y}(r)$ occurs when $\mathbf{Y}(r) = \tilde{\mathbf{X}}$. This is attained when $r = r^*$ is chosen, meaning that the value of the expression at its minimum is $-\frac{1}{2} \tilde{\mathbf{X}}^T \mathbf{D} \tilde{\mathbf{X}} - H$. Finally, since the true signal $\tilde{X}$ is unknown, this value is unknown. However, since it is known that $\tilde{X}_l = Y_l(r^*)$, it can be said that at minimum, it will be

$$-\frac{1}{2} \tilde{\mathbf{X}}^T \mathbf{D} \tilde{\mathbf{X}} \geq \min_r -\frac{1}{2} \mathbf{Y}(r)^T \mathbf{D} \mathbf{Y}(r)$$

attained at $\hat{r}$ with $\hat{\mathbf{Y}} = \mathbf{Y}(\hat{r})$. This is a value that can be easily computed for the given model M , since it is independent of the image X and also doesn't depend on poses or shifts r, t . It is the CTF-corrupted power at high frequencies of the slice of M that has the most power.

[0103] Thus, it is finally arrived at

$$V_2 - V_3 \geq -\sum_{\|l\| > L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2 - H$$

[0104] From the above, it follows that:

$$E(r, t) \geq U(r, t) + V_1 - \sum_{\|l\| > L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2 - H$$

[0105] Due to the presence of H , this expression provides a probabilistic bound on E , indicating that the probability of E is greater than the value of the expression. In an example, a probability, 0.999936, which corresponds to four standard deviations of H , is selected to arrive finally at the lower bound on $E(r, t)$:

$$E(r, t) \geq U(r, t) + V_1 - \sum_{\|l\| > L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2 - 4\sigma_H \quad (19)$$

$$= U(r, t) + V_1 - \sum_{\|l\| > L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2 - 4\sqrt{\sum_{\|l\| > L} \frac{\sigma^2}{2} C_l^2 |Y_l(r)|^2} \quad (20)$$

$$\geq U(r, t) + V_1 - \sum_{\|l\| > L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2 - 4 \sqrt{\sum_{\|l\| > L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2} \quad (21)$$

$$= B_L(r, t) \quad (22)$$

In the last line, the maximum power slice $\hat{Y}$ is again used, and that $\sigma = 1$.

[0106] At this point, an expression is arrived at which, with very high probability, bounds $E(r, t)$ from below. This expression is cheap to compute for an particular r, t (since only U depends on these), and the remainder only needs to be computed once for all r, t , and also only once for all images that share the same CTF; i.e., that come from the same micrograph. To compute the bound, it is only needed to find the slice of the model which has the most power ($\hat{Y}$). Then the expression is used to compute the values of U and the power of the image V_1 .

[0107] The bound above is actually a family of bounds, one for each radius L . As L is increased, the bound becomes more expensive but tighter, and finally when L reaches the Nyquist rate then the bound becomes exact, but is as expensive as directly computing $E(r, t)$.

[0108] Returning to Fig. 2, upon establishing a lower bound, an upper bound is established (115).

[0109] As an upper bound on $E(r, t)$, a single direct evaluation of E at a specific pose and shift can be used. During optimization, a candidate best pose and shift r^*, t^* are maintained, and are used to compute $E^* = E(r^*, t^*)$, the upper bound on the minimum of E . It can be said that the minimum of $E(r, t)$ can not be greater than E^* , because that would be a contradiction - every $E(r, t)$ for all r, t must be greater or equal to the value at the minimum.

[0110] Fig. 5A shows an image of a protein taken with an electron microscope. Figs. 5B to 5D show the image of 5A after application of various filters.

[0111] Figs. 6 to 8 show values of the lower bound $B_L(r, t)$ at various levels L for the image of Fig. 5A showing lower and upper bound being useful at different radii. In each plot, the black horizontal line is an upper bound E^* , while the other lines are the lower bounds. After evaluating a single bound, say B_{12} , only the regions below the horizontal black line need to be re-evaluated at a higher level, say B_{30} . The true minimum of E must be between the upper and lower bounds. Fig. 6 shows the values of a bound over poses r , taking the minimum over all shifts to arrive at the values for each r . Each position on the x-axis of the plot corresponds to the index of a unique pose r on a grid covering all possible poses. Fig. 7 shows a zoomed in section of the graph of Fig. 6 near the true best pose. Note that the upper bound intersects the lower bounds,

1 meaning that evaluating a lower bound (rather than the full objective $E(r, t)$) allows one to
 2 discard vast regions of the pose space. Also note that the family of bounds B_L becomes tighter
 3 (higher) with increasing L , until finally the upper and lower bounds meet at the true minimum of
 4 the objective. Fig. 8 shows the same bounds, but over shifts t rather than poses. Each position
 5 on the x-axis corresponds to a unique position on a grid covering a range of possible shifts.

6 [0112] The lower bound derived above, $B_L(r, t)$, can be used to develop a method to search
 7 over poses and shifts faster than exhaustive search without loss of alignment accuracy. With the
 8 theoretical support of the bound in place, very similar but more effective approximate bounds
 9 and techniques are described.

10 [0113] Referring again to Fig. 2, from the starting point of exhaustive search, the following
 11 BnB optimization using the bound B_L is introduced. Here $\mathbf{R}_0$ and $\mathbf{T}_0$ are the discrete grids that
 12 would have been used for exhaustive search. L is initialized as $L = L_0$ (120). $\mathbf{R}$ and $\mathbf{T}$ are
 13 initialized as $\mathbf{R} = \mathbf{R}_0, \mathbf{T} = \mathbf{T}_0$ (130). While it is determined at 140 that a selected accuracy of
 14 image alignment has been obtained, in this case being that L is less than the Nyquist rate,
 15 poses are evaluated (150), after which L is doubled (155). In further cases, the selected
 16 accuracy of image alignment has been obtained when a selected number of iterations have been
 17 performed.

18 [0114] Fig. 9 shows the process of evaluating poses at 150. $B_L(r, t)$ is computed for all r, t
 19 in $\mathbf{R}, \mathbf{T}$ (151). The r^*, t^* with minimum B_L are located (152). $E^* = E(r^*, t^*)$ is evaluated (153).
 20 For every r in $\mathbf{R}$, if $\min_t B_L(r, t)$ is greater than E^* , r is removed from $\mathbf{R}$ (154). For every t
 21 in $\mathbf{T}$, if $\min_r B_L(r, t)$ is greater than E^* , t is removed from $\mathbf{T}$ (155).

22 [0115] Returning again to Fig. 2, once $L \geq N/2$, r^*, t^* are returned, which represent the pose
 23 and shift of the optimal alignment of the image to the structure.

24 [0116] As this technique progresses, poses and shifts are identified that cannot be the
 25 minimum E because the lower bound at that pose is greater than the current best value of E . As
 26 these points are identified, they are removed from the set of candidate poses and shifts. Upon
 27 termination, only the pose and shift with minimum E will remain.

28 [0117] Note that this method begins by evaluating $B_{L_0}(r, t)$ for all of $\mathbf{R}, \mathbf{T}$, as opposed to
 29 exhaustive scan, which would have evaluated $E(r, t)$ for all $\mathbf{R}, \mathbf{T}$. The cost of computing $E(r, t)$
 30 is proportional to $(N/2)^2$ and the cost of computing $B_L(r, t)$ is proportional to L^2 . This means

1 that, at best, the speedup that can be achieved with this technique is $(\frac{N/2}{L_0})^2$. In some scenarios
 2 this is about 100, but speedups on the order of 20-30 can be achieved in other scenarios.

3 [0118] The bound described above gives a strong theoretical backing for the assumption, and
 4 also allows modest speedups to be achieved.

5 First Level Approximations

6 [0119] In another implementation, some approximations can be made to yield a method that
 7 is similar, but provides much larger speedups. The existence and tightness of bound $B_L(r, t)$
 8 indicate that low-resolution alignment provides significant information about high-resolution
 9 alignment.

10 [0120] Consider first bound B_L . To compute it, the slice Y of model M which has the
 11 largest CTF-corrupted power can be found. One drawback of this computation is that it needs to
 12 be redone for each set of images that has a different CTF. The CTF, at frequencies above L , is
 13 like a sine function with a varying frequency. The RMS power of the CTF is thus close to $1/\sqrt{2}$.
 14 An approximation can be made and it can be assumed that

$$15 \quad \sum_{\|l\|>L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2 \approx \frac{1}{2} \sum_{\|l\|>L} \left(\frac{1}{\sqrt{2}}\right)^2 |\hat{Y}_l|^2 = \sum_{\|l\|>L} \frac{1}{4} |\hat{Y}_l|^2$$

16 [0121] This expression no longer depends on the CTF, so the approximate bound can be
 17 entirely independent of the image, and only needs to be computed once, given the model. Note
 18 that this approximation of the CTF does not actually make the bound tighter, rather it removes
 19 the dependence on the image.

20 [0122] Again consider the above expression. A search is performed over slices Y to find the
 21 one with the maximum power. The maximum power slice, $\hat{Y}$, gives a strict limit on how much
 22 power there could be at high frequencies that might contribute to E . However, it is very unlikely
 23 that for any given image, the true pose will be the same as the pose of $\hat{Y}$, since most images do
 24 not come from the maximum power slice. Thus, the slice $Y^\dagger$ that has the median power of all
 25 slices can be used as an approximation instead of the maximum (or any other rank statistic).

26 [0123] All together, the above approximations lead to an approximate lower bound for
 27 $E(r, t)$ as

$$28 \quad A_L(r, t) = U(r, t) + V_1 - \sum_{\|l\|>L} \frac{1}{4} |Y_l^\dagger|^2 - 4 \sqrt{\sum_{\|l\|>L} \frac{1}{4} |Y_l^\dagger|^2}$$

1 [0124] So far, the approximations allow for the construction of an approximate bound that is
 2 inexpensive, but this is still limited in the same way as the application of the true bound in
 3 speeding up exhaustive scan (previous implementation). To overcome this limitation, the
 4 following assumption is applied: when the bound at low resolutions (low L) is evaluated, the
 5 sampling need not be nearly as fine in pose and shift space as needed when searching at high
 6 resolution. At low L , coarse sampling along with the bounds themselves can allow the
 7 identification of regions that need to be searched further at high L and finer sampling.

8 [0125] One way to understand why a coarse (at low resolution) to fine (at high resolution)
 9 sampling approach is reasonable is to consider the angular spacing of Fourier coefficients in
 10 each Fourier shell of a 3D structure. This spacing is much larger at low resolutions than at high
 11 resolutions.

12 [0126] A stronger argument can be made by considering the fact that the 3D structure in real
 13 space is compact and has a small extent relative to the full volume. This is true because of the
 14 masking that is applied during refinement of a 3D structure for Cryo-EM. Usually, in Fourier
 15 analysis, if the Fourier transform $F(i, j, k)$ of a real-space signal $f(x, y, z)$ has no power a
 16 distance d away from the origin of Fourier space, it can be said that the real-space f is
 17 'band-limited', has no high-frequency power, and is thus smooth. The Fourier transform,
 18 however, is a unitary transform, so this argument can be exactly inverted. It can be said that if a
 19 real-signal f has no power a distance d away from the origin of real space, then the Fourier
 20 transform F must be smooth. This reasoning indicates that $B_L(r, t)$, which depends on slices of
 21 the 3D Fourier transform of a masked structure, is smooth with respect to r, t . The smoothness
 22 is strongest when L is small.

23 [0127] A technique can be used to exploit this notion of smoothness. The pose and shift
 24 space are first subdivided into cells on a coarse grid R_0, T_0 . On this grid, the (approximate)
 25 bound is evaluated at a low value of L . This allows for the identification of cells that may contain
 26 the minimum E . All those cells are kept as candidates. Each candidate cell is then subdivided by
 27 a factor of 2 in each dimension, increase L , and repeat the process.

28 [0128] Domain subdivision of this type removes the speedup limitation from the previous
 29 sections. With domain subdivision, approximate branch and bound alignment can reach several
 30 orders of magnitude of speedup relative to exhaustive scan.

31 [0129] Together, the approximate bounds and subdivision scheme allow for the construction
 32 of an effective technique for alignment.

[0130] The process of evaluating poses in this implementation is shown generally at 200 in Fig. 10. $A_L(r, t)$ is computed for all r, t in R, T (approximate lower bound) (201). The r^*, t^* with minimum A_L are then located (202). E^* is set to $E^* = E(r^*, t^*)$ (203). For every r in R , if $\min_t A_L(r, t)$ is greater than E^* , r is removed from R (204). For every t in T , if $\min_r A_L(r, t)$ is greater than E^* , t is removed from T (205). Replace every r in R with 8 subdivided points (206). Replace every t in T with 4 subdivided points (207).

[0131] This process in this second implementation works very well on real data, achieving significant speedups in many cases. There is one tuning parameter, the coarseness of the initial grids R_0 and T_0 , that allows a trade-off between accuracy and speed. On real images, even liberal settings of this parameter do not affect the quality of alignments, but give significant speedups.

12 Second Level Approximations

[0132] In a further implementation, a further simplified approximate version of the BnB optimization is developed that can be implemented very simply. Empirically, this further simplified approximate version produces alignments of high quality on real data, resulting in high resolution 3D structures, while being extremely fast.

[0133] In practice, the bound and approximate bound presented in earlier sections reject large regions of the pose space, for most images. On some images, however, the bounds do not reject enough of the pose space at each iteration of the alignment technique, meaning that the method does not provide much speedup on those images.

[0134] On further observation, for a typical dataset, it is only a small fraction of images that perform poorly, and that these images are actually outlier images, meaning that they are not images of the structure to which alignments are computed. They may be images that were incorrectly picked from micrographs, or they may be images of broken/damaged particles. The characteristic of these images is that they break the assumptions that make the bound tight (and thus fast). These assumptions are that the observed image is a noisy slice from the model M , and that the image has the correct formation and noise model. In a following section, it is discussed how to directly detect the breaking of these assumptions and thus reject outlier images outright.

[0135] To deal with the problem of some pathological images in alignment, the upper bound is replaced with a fixed fraction f_{keep} . At each iteration of the alignment technique, only the best f_{keep} of the candidate poses and shifts are kept, discarding the rest. This effectively means that

1 all images will take the same amount of computation to align, and the value of f_{keep} will be
 2 another trade-off parameter between speed and accuracy. Using this modification to the
 3 alignment technique eliminates the need for even computing the upper or lower bounds.

4 [0136] The fraction of poses and shifts that are discarded on typical images on a typical
 5 dataset using the approximate bound are empirically examined, and f_{keep} is set accordingly.
 6 The Applicant has determined that, in some cases, a fixed fraction f_{keep} of between 3% and
 7 10% is effective. In a further case, the Applicant has determined that a fixed fraction f_{keep} of
 8 approximately 5% is effective.

9 [0137] The process of evaluating poses in this implementation is shown generally at 300 in
 10 Fig. 11. $U(r, t)$ is computed for all r, t in R, T (301). U^* is set so that $\frac{|{(r,t); U(r,t) \leq U^*}|}{|{(r,t); U(r,t) > U^*}|} = f_{\text{keep}}$
 11 (302). The r^*, t^* with minimum $U(r, t)$ is located (303). For every r in R , if $\min_t U(r, t)$ is
 12 greater than U^* , r is removed from R (304). For every t in T , if $\min_r U(r, t)$ is greater than
 13 U^* , remove t from T (305). Replace every r in R with 8 subdivided points (306). Replace
 14 every t in T with 4 subdivided points (307).

15 [0138] Although this implementation does not guarantee optimal alignment, it is an
 16 approximation of the mathematically sound technique and bound proven to be correct earlier.
 17 This, along with excellent empirical performance, sets it apart from existing heuristics.

18 [0139] To demonstrate the empirical performance of the alignment technique using second
 19 level approximations, 3D structures reconstructed using the technique on two different datasets
 20 are shown, together with the computational time it took to recover these structures.

21 [0140] Figs. 12A to 12C show three slices from the x, y, and z directions of a result of the
 22 determination of the three-dimensional structure of the TRPV1 protein using the alignment
 23 technique presented herein. Fig. 13 is an Fourier Shell Correlation (FSC) resolution estimate of
 24 the result in Figs. 12A to 12C. Fig. 14 is an image that illustrates the quality of the structure
 25 reconstructed using the process of the third implementation, with side-chains visible in alpha
 26 helices. This structure was computed in less than one hour on a single desktop
 27 GPU-workstation.

28 [0141] Figs. 15A to 15C show three slices from the x, y, and z directions of a result of the
 29 determination of the three-dimensional structure of the 80S ribosome protein using the alignment
 30 technique presented herein. Fig. 16 shows an FSC resolution estimate of the result in Figs. 15A

1 to 15C. This structure was computed again in less than two hours on a single desktop
2 GPU-workstation.

3 Representation, Discretization and Subdivision of Pose (r) and Shift (t)

4 [0142] In the present embodiments, poses and shifts are represented by numerical vector
5 quantities. The present embodiments apply three-dimensional vectors to encode poses using the
6 axis-angle formulation; however, it is appreciated that poses can be parameterized in any
7 suitable fashion. The present embodiments parameterize shifts by applying two-dimensional
8 vectors with units of pixels of shift in the horizontal and vertical directions; however, it is
9 appreciated that shifts can be parameterized in any suitable fashion.

10 [0143] The space of poses, as vectors in the axis-angle formulation, comprises a sphere of
11 radius π centered at the origin. This space can be discretized into a grid of possible poses where
12 each pose is a three-dimensional grid point. The grid can be instantiated with a desired fineness,
13 to cover the sphere of radius π . Once a grid is instantiated with a grid spacing of d , it can be
14 subdivided into a grid of finer coarseness by splitting the grid in each dimension into grid cells of
15 size $d/2$. This yields eight new subdivided grid cells for each grid cell in the original grid.

16 [0144] Similarly, the space of shifts of an image can be discretized into a two dimensional
17 grid of desired fineness covering a desired range of image shifts. This two dimensional grid of
18 grid spacing d can be subdivided by splitting each grid cell into four cells of size $d/2$.

19 [0145] The representations, discretizations and subdivision methods, described herein, can
20 also be applied when the pose space is one-dimensional or when the shift space is three
21 dimensional, as may be the case in other embodiments.

22 Termination Criteria

23 [0146] In the above-described implementations, the coarseness of the initial pose and shift
24 grids determines the precision that is reached after K iterations of subdivision. In practice,
25 alignment iterations are not terminated when $L \geq N/2$, but rather continue until the precision of
26 alignment reaches a reasonable level. This level can be determined using a common measure of
27 how well the grid is sampling $E(r, t)$, known as the Effective Sample Size:

$$28 \quad e = \frac{1}{\sum_{r,t} \left(\frac{\exp(-E(r,t))}{\sum_{r,t} \exp(-E(r,t))} \right)^2}$$

29 [0147] This value e represents the number of poses and shifts that have significant
30 probability of being the correct pose. On a coarse grid, there may only be one pose or shift that

1 has significant probability. As the pose and shift grids are subdivided, e increases. Once e
 2 reaches a minimum value $e_{\min}$ for the majority of images being aligned (i.e. the median e over
 3 images), alignment can be terminated. Typical values are $2 \leq e_{\min} \leq 20$.

4 Mode Marginalization

5 [0148] The above-described implementations can be used to find the best looking pose and
 6 shift for an image and reference structure. However, in practice, it is common to marginalize over
 7 all poses and shifts, averaging together their contributions weighted by the probability of each
 8 pose and shift. Typically only a small number of poses and shifts have significant probability, and
 9 they are all concentrated in a mode around the maximally probable r, t .

10 [0149] To support marginalization, the maximally probable pose is first found using the
 11 alignment techniques presented in this work. Marginalization is subsequently performed only
 12 over poses near to this pose, which provides the benefits of marginalization without the extreme
 13 computational cost of current exhaustive approaches.

14 Use in 2D/2D and 3D/3D alignment

15 [0150] The implementations described above are directed to the alignment of 3D reference
 16 structure to a set of 2D images. The approach taught herein, however, can be equally useful when
 17 aligning a set of 2D images to a 2D reference. This is often the case in what is known as class
 18 averaging or 2D classification. The methods of all the implementations above can be used for
 19 2D-to-2D alignment simply by applying them over a pose space that contains only one dimension
 20 of in-plane rotation rather than three dimensions of 3D pose.

21 [0151] Similarly, the methods of all the implementations above can be used for 3D-to-3D
 22 alignment simply by applying them over a three dimensional pose space as described above, but
 23 applying them over a three dimensional shift space rather than the two dimensions described
 24 above.

25 Use in heterogeneous reconstruction

26 [0152] The implementations described above are directed to reconstructing a single 3D
 27 structure from a set of 2D images. The approach taught herein, however, can be equally useful
 28 when reconstructing multiple structure from a set of images. This is known as heterogeneous
 29 reconstruction, or 3D classification. The goal is to separate images into classes, each with a
 30 corresponding 3D structure, in order to resolve multiple different structures that may have been
 31 simultaneously present in a sample during imaging.

1 [0153] The methods of all of the implementations above can be used for heterogeneous
2 reconstruction by applying them individually to each structure class, and assigning (potentially
3 using soft weights rather than hard assignment) each image to the classes based on the final
4 error $E(r, t)$ for each class. An image will belong with more weight to the classes with lower
5 error.

6 Use for added degrees of freedom (advanced heterogeneity, flexibility)

7 [0154] Similarly to heterogeneous reconstruction, where an additional degree of freedom is
8 introduced into the problem in the form of class assignment, the methods of the implementations
9 described above can be used in more advanced types of heterogeneity like combinatorial
10 heterogeneity where each image is modeled as a combination of a set of structures, each
11 structure having a binary presence variable. The techniques can also be used when attempting
12 to deal with flexible structures, for instance by parameterizing the flexibility using additional
13 degrees of freedom.

14 [0155] In these cases, the methods of the implementations can be used directly to compute
15 a minimum over the new error function $E(r, t, v)$ where v are the extra degrees of freedom
16 introduced. This can be done by brute force; i.e., by applying the techniques as they are to the
17 new error function, fixing v and repeating for different values of v on a grid over the search
18 space of v . Alternatively the methods of the implementations can be extended by bounding and
19 subdividing over v as well as r and t .

20 [0156] While the above-described embodiments have been described with respect to
21 traditional defocus images from transmission electron microscopes, they can also be applied to
22 phase plates for these microscopes without any modification, requiring only that the correct CTF
23 model be used. The above-described approaches perform well on phase plate data due to the
24 added power at low frequencies.

25 [0157] Although the invention has been described with reference to certain specific
26 embodiments, various modifications thereof will be apparent to those skilled in the art without
27 departing from the spirit and scope of the invention as outlined in the claims appended hereto.

CLAIMS

We claim:

1. A method for image alignment of at least one two-dimensional or three-dimensional image to a two-dimensional or three-dimensional model, executed on a processing unit, the image alignment having an acceptable likelihood of mismatch between the at least one image and the model, the method comprising:

selecting a value for a radius in Fourier space;

discretizing a set of poses into a discrete grid of candidate poses and a set of shifts into a discrete grid of candidate shifts;

determining a fixed fraction (f_{keep}) as an upper bound on the acceptable likelihood of mismatch, the fixed fraction being determined based on, at least, the fraction of poses and shifts that are discarded on typical dataset of images;

determining whether a selected accuracy of image alignment has been obtained, when such determination is false:

using the upper bound and a lower bound on the acceptable likelihood of mismatch, the lower bound being a combination of a first component ($U(r, t)$) and a second component ($V(r, t)$), the second component comprising a first subcomponent added to a second subcomponent and subtracted by a third subcomponent, the first subcomponent being the power of one of the images at high frequencies, the second subcomponent being the power of an image of a slice of the three-dimensional model for one of the poses at high frequencies, the third subcomponent being the correlation between the power of each of the images at high frequencies and the power of the image of the slice of the three-dimensional model for one of the poses at high frequencies;

determining values for the first component for each of the poses and shifts on the discrete grid of candidate poses and the discrete grid of shifts, using only portions of the image below the selected value for the radius in Fourier space;

determining a reference first component (U^*) using the fixed fraction;

for every pose in the discrete grid of candidate poses, determining whether a minimum value of the first component over all the candidate shifts is greater than the reference first component and discarding the pose from the discrete grid in such case;

for every shift in the discrete grid of candidate shifts, determining whether the minimum value of the first component over all candidate poses is greater than the reference first component and discarding the shift from the discrete grid in such case;

for every remaining pose in the discrete grid of candidate poses, replacing the pose with a plurality of subdivided grid points representing the candidate poses;

for every remaining shift in the discrete grid of candidate shifts, replacing the shift with a plurality of subdivided grid points representing the candidate shifts; and

increasing the radius in the Fourier space; and

otherwise, returning the pose and shift at the lower bound with minimum value.

2. The method of claim 1, wherein the lower bound is determined with the images at a resolution that is less than the maximum resolution for the images.
3. The method of claims 1 or 2, wherein the first component is the squared error of Fourier coefficients at or below a selected radius in Fourier space, and the second component is the squared error of Fourier coefficients above the selected radius.
4. The method of any one of claims 1 to 3, wherein the second subcomponent comprises:

$$V_1 - \sum_{\|l\| > L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2 - 4 \sqrt{\sum_{\|l\| > L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2} ,$$

wherein V_1 is the power of one of the images at high frequencies, subscript l denotes a wavevector, subscript L denotes the selected radius in the Fourier space, C is a contrast transfer function (CTF), and Y is a vector representing a projection of the three-dimensional model.

5. The method of claim 4, wherein the second subcomponent is only recomputed if the contrast transfer function is different.

6. The method of any one of claims 1 to 5, wherein the determination of the upper bound comprises evaluating a value for the likelihood of mismatch at a specific pose, specific shift, or both.
7. The method of any one of claims 1 to 6, wherein determining the reference first component (U^*) comprises determining the reference first component (U^*) such that:

$$\frac{|\{(r,t); U(r,t) \leq U^*\}|}{|\{(r,t); U(r,t) > U^*\}|} = f_{\text{keep}} .$$

8. The method of any one of claims 1 to 7, wherein the at least one image is two-dimensional and the model is three-dimensional, wherein replacing the pose with the plurality of subdivided grid points comprises replacing the pose with eight subdivided grid points, and wherein replacing the shift with the plurality of subdivided grid points comprises replacing the shift with four subdivided grid points.
9. The method of any one of claims 1 to 7, wherein the at least one image is two-dimensional and the model is two-dimensional, wherein replacing the pose with the plurality of subdivided grid points comprises replacing the pose with two subdivided grid points, and wherein replacing the shift with the plurality of subdivided grid points comprises replacing the shift with four subdivided grid points.
10. The method of any one of claims 1 to 7, wherein the at least one image is three-dimensional and the model is three-dimensional, wherein replacing the pose with the plurality of subdivided grid points comprises replacing the pose with eight subdivided grid points, and wherein replacing the shift with the plurality of subdivided grid points comprises replacing the shift with eight subdivided grid points.
11. The method of any one of claims 1 to 10, wherein the selected accuracy of image alignment is obtained when the value for the radius in Fourier space is equal to the Nyquist rate.
12. The method of any one of claims 1 to 10, wherein the selected accuracy of image alignment is obtained when a selected number of iterations have been performed.
13. The method of any one of claims 1 to 12, wherein the fixed fraction (f_{keep}) is between 3% and 10%.
14. The method of any one of claims 1 to 13, wherein the fixed fraction (f_{keep}) is 5%.
15. A method for image alignment of at least one two-dimensional or three-dimensional image to a two-dimensional or three-dimensional model, executed on a processing unit, the image

alignment having an acceptable likelihood of mismatch between the at least one image and the model, the method comprising:

- selecting a value for a radius in Fourier space;
- discretizing a set of poses into a discrete grid of candidate poses and a set of shifts into a discrete grid of candidate shifts;
- determining a fixed fraction as an upper bound on the acceptable likelihood of mismatch, the fixed fraction being determined based on, at least, the fraction of poses and shifts that are discarded on typical dataset of images;
- the processing unit iteratively determining whether a selected accuracy of image alignment has been obtained, and when such determination is false:
 - assigning a lower bound to the acceptable likelihood of mismatch, the lower bound comprising a first component;
 - analyzing the image to isolate selected portions of the image that are below the value for a radius in Fourier space;
 - determining values for the first component for each of the poses and shifts on the discrete grid of candidate poses and the discrete grid of shifts, using only the isolated selected portions of the image;
 - determining a reference first component using a value for the fixed fraction;
 - parsing the discrete grid of candidate poses by analyzing each one of the candidate poses over all of the candidate shifts to obtain a minimum value of the first component over all the candidate shifts and discarding each candidate pose from the discrete grid if the first component exceeds the reference first component;
 - parsing the discrete grid of candidate shifts by analyzing each of the candidate shifts over all of the candidate poses to obtain a minimum value of the first component over all candidate poses and discarding each candidate shift from the discrete grid if the first component exceeds the reference first component;
 - for every remaining pose in the discrete grid of candidate poses, replacing the pose with a plurality of subdivided grid points representing the candidate poses;

for every remaining shift in the discrete grid of candidate shifts, replacing the shift with a plurality of subdivided grid points representing the candidate shifts; and

increasing the radius in the Fourier space; and

otherwise, returning the pose and shift at the lower bound with minimum value.

16. The method of claim 15, wherein the lower bound is determined with the images at a resolution that is less than the maximum resolution for the images.
17. The method of claims 15 or 16, wherein the first component is the squared error of Fourier coefficients at or below a selected radius in Fourier space, and wherein the lower bound further comprises a second component that is the squared error of Fourier coefficients above the selected radius.
18. The method of any one of claims 15 to 17, wherein the lower bound further comprises a second component that comprises:

$$V_1 - \sum_{\|l\| > L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2 - 4 \sqrt{\sum_{\|l\| > L} \frac{1}{2} C_l^2 |\hat{Y}_l|^2} ,$$

wherein V_1 is the power of one of the images at high frequencies, subscript l denotes a wavevector, subscript L denotes the selected radius in the Fourier space, C is a contrast transfer function (CTF), and Y is a vector representing a projection of the three-dimensional model.

19. The method of claims 17 or 18, wherein a subcomponent of the second component is only recomputed if the contrast transfer function is different.
20. The method of any one of claims 15 to 19, wherein the determination of the upper bound comprises evaluating a value for the likelihood of mismatch at a specific pose, specific shift, or both.
21. The method of any one of claims 15 to 20, wherein determining the reference first component (U^*) comprises determining the reference first component (U^*) such that:

$$\frac{|\{(r,t); U(r,t) \leq U^*\}|}{|\{(r,t); U(r,t) > U^*\}|} = f_{\text{keep}} , \text{ wherein } f_{\text{keep}} \text{ is the fixed fraction.}$$

22. The method of any one of claims 15 to 21, wherein the at least one image is two-dimensional and the model is three-dimensional, wherein replacing the pose with the plurality of subdivided grid points comprises replacing the pose with eight subdivided grid points, and wherein

replacing the shift with the plurality of subdivided grid points comprises replacing the shift with four subdivided grid points.

23. The method of any one of claims 15 to 21, wherein the at least one image is two-dimensional and the model is two-dimensional, wherein replacing the pose with the plurality of subdivided grid points comprises replacing the pose with two subdivided grid points, and wherein replacing the shift with the plurality of subdivided grid points comprises replacing the shift with four subdivided grid points.
24. The method of any one of claims 15 to 21, wherein the at least one image is three-dimensional and the model is three-dimensional, wherein replacing the pose with the plurality of subdivided grid points comprises replacing the pose with eight subdivided grid points, and wherein replacing the shift with the plurality of subdivided grid points comprises replacing the shift with eight subdivided grid points.
25. The method of any one of claims 15 to 24, wherein the selected accuracy of image alignment is obtained when the value for the radius in Fourier space is equal to the Nyquist rate.
26. The method of any one of claims 15 to 24, wherein the selected accuracy of image alignment is obtained when a selected number of iterations have been performed.
27. The method of any one of claims 15 to 26, wherein the fixed fraction is between 3% and 10%.
28. The method of any one of claims 15 to 27, wherein the fixed fraction is 5%.

1 / 18

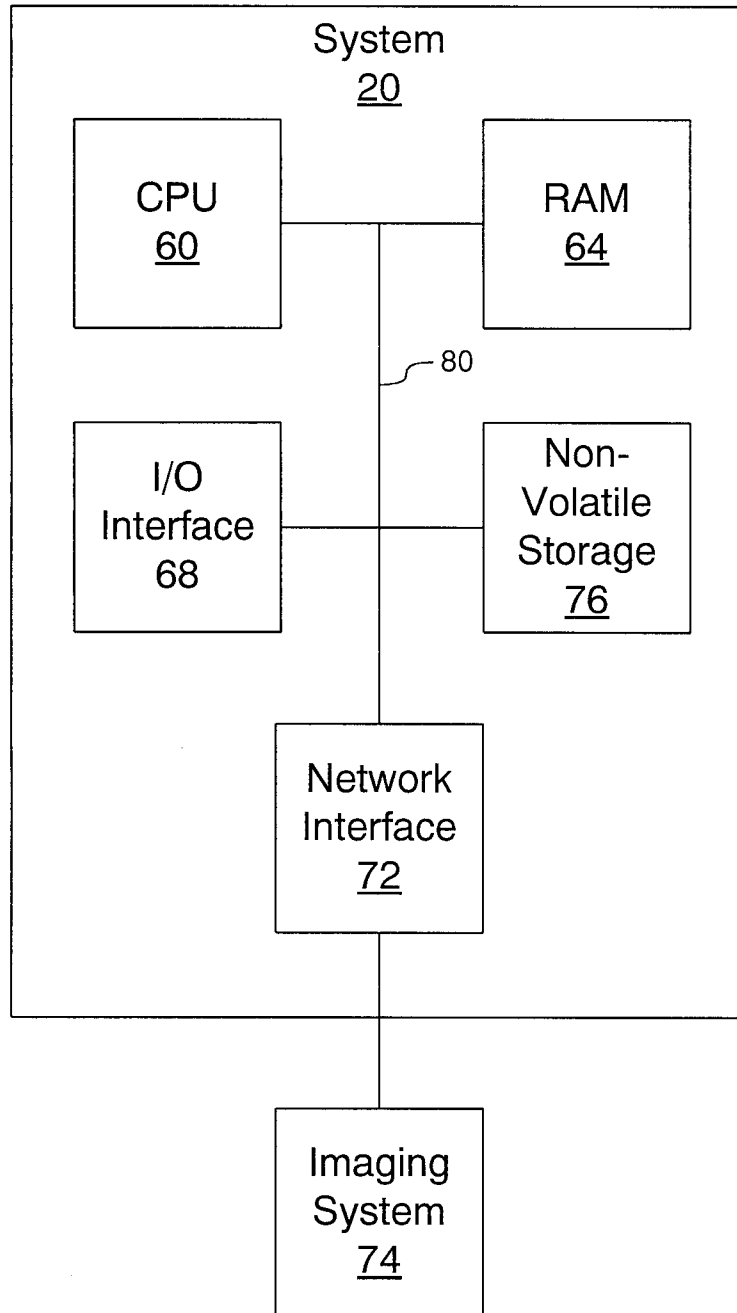


FIG. 1

2 / 18

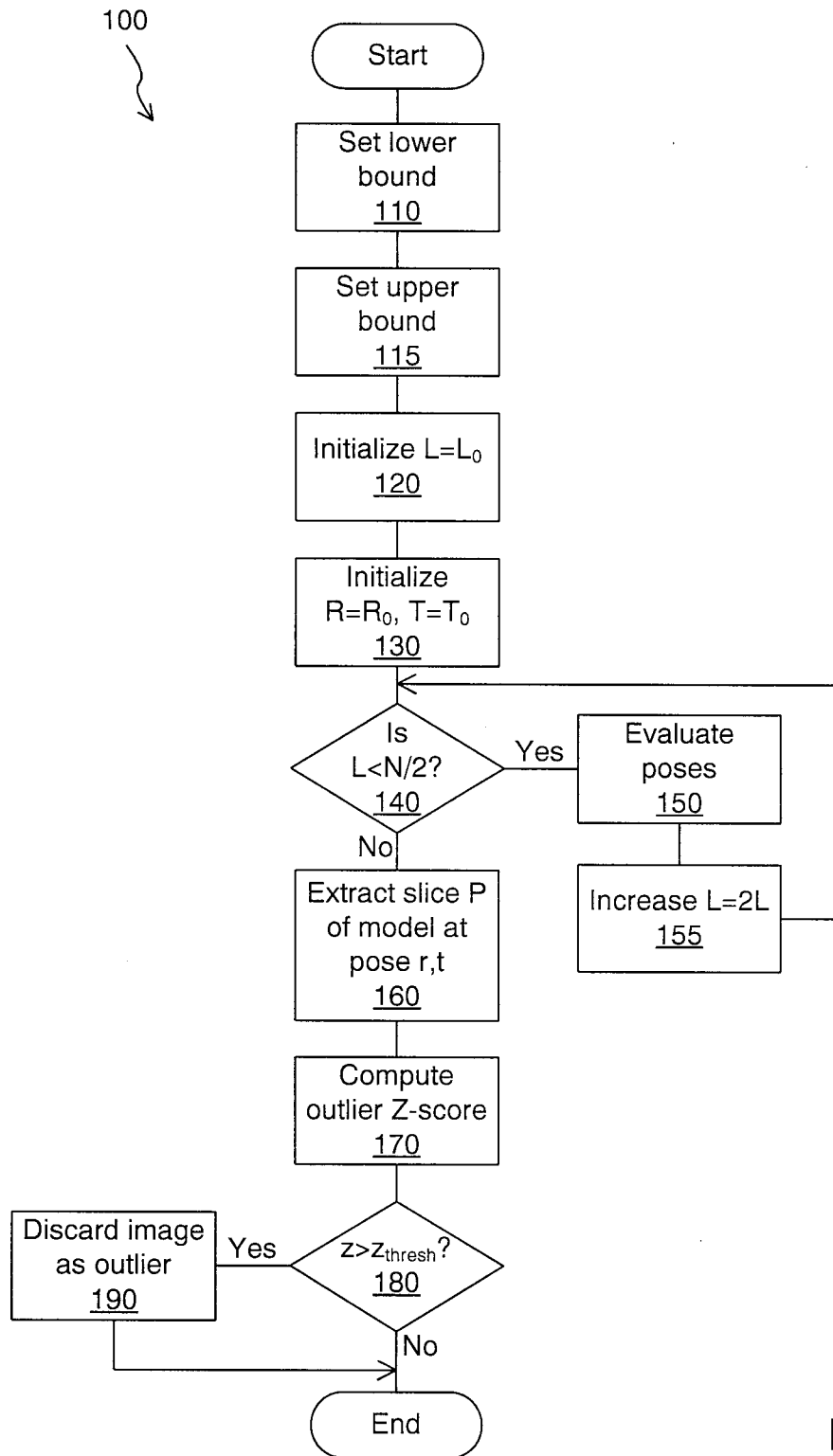


FIG. 2

3 / 18

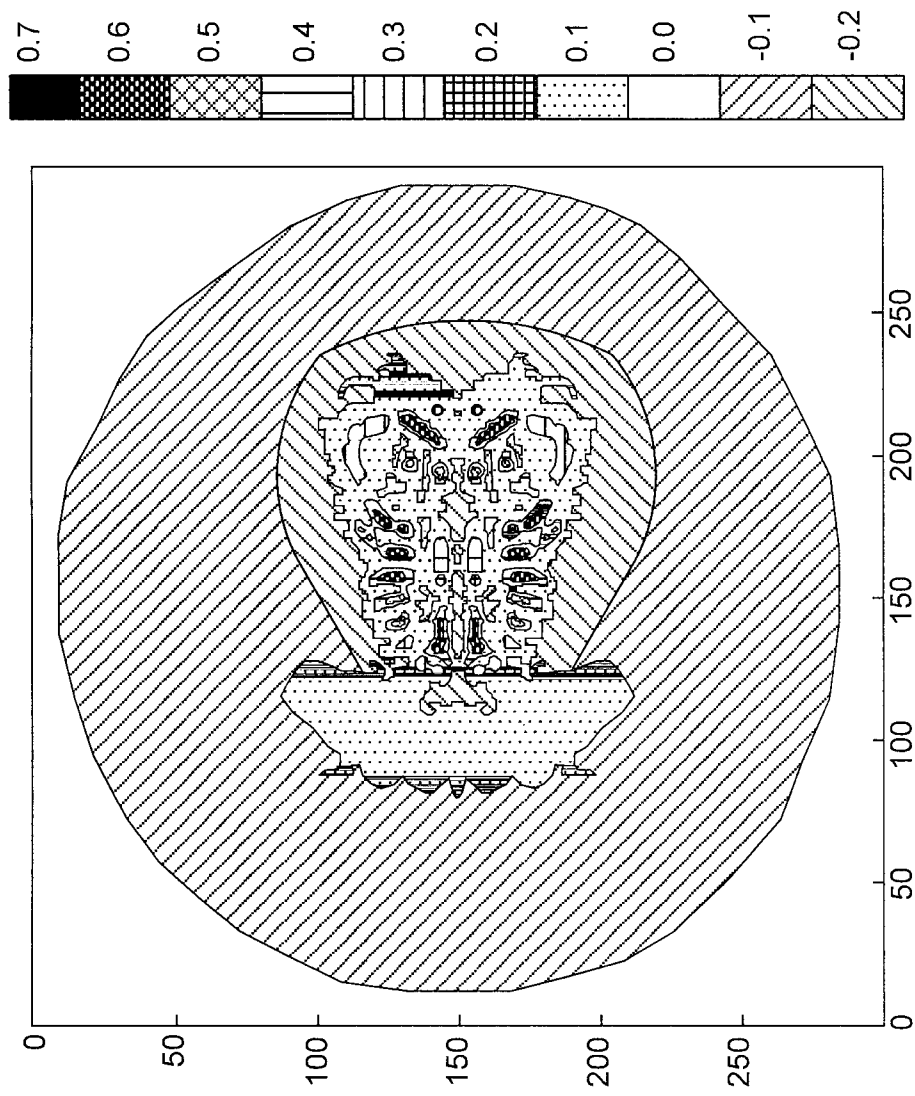


FIG. 3A

4 / 18

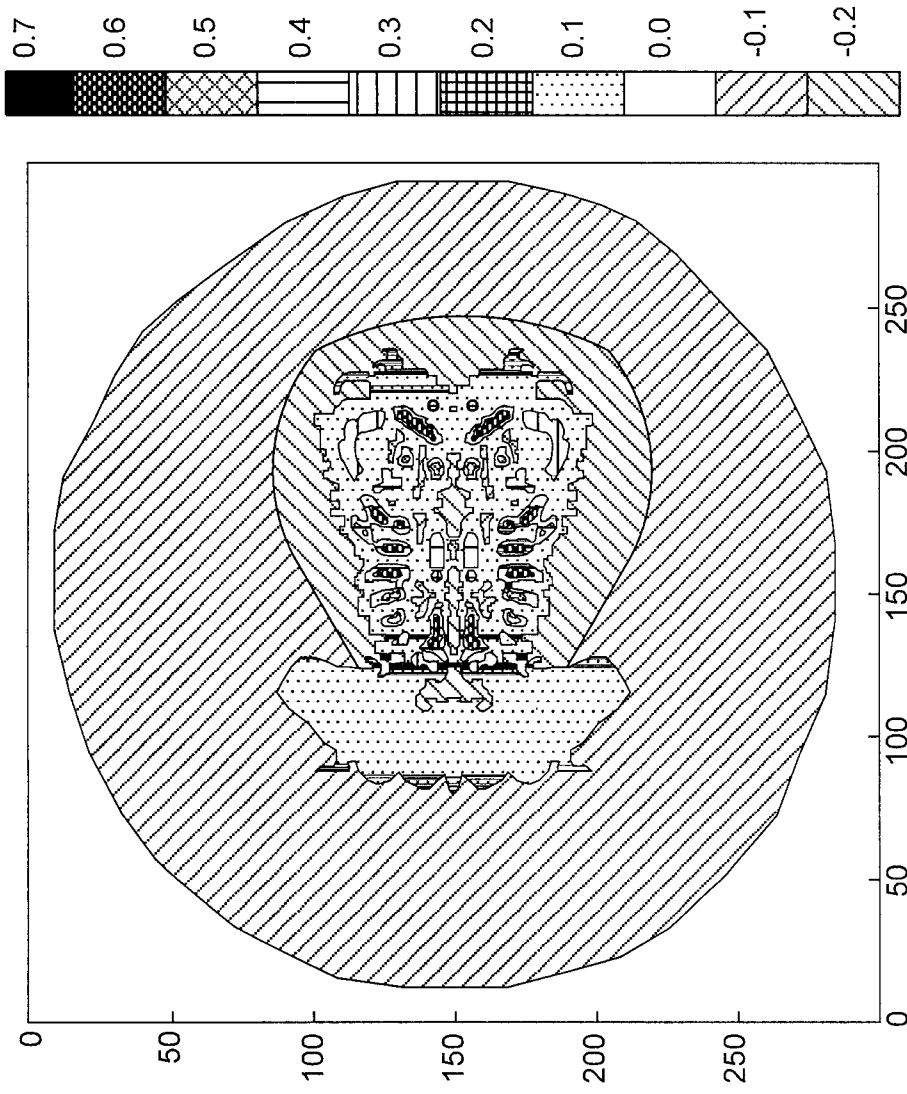


FIG. 3B

5 / 18

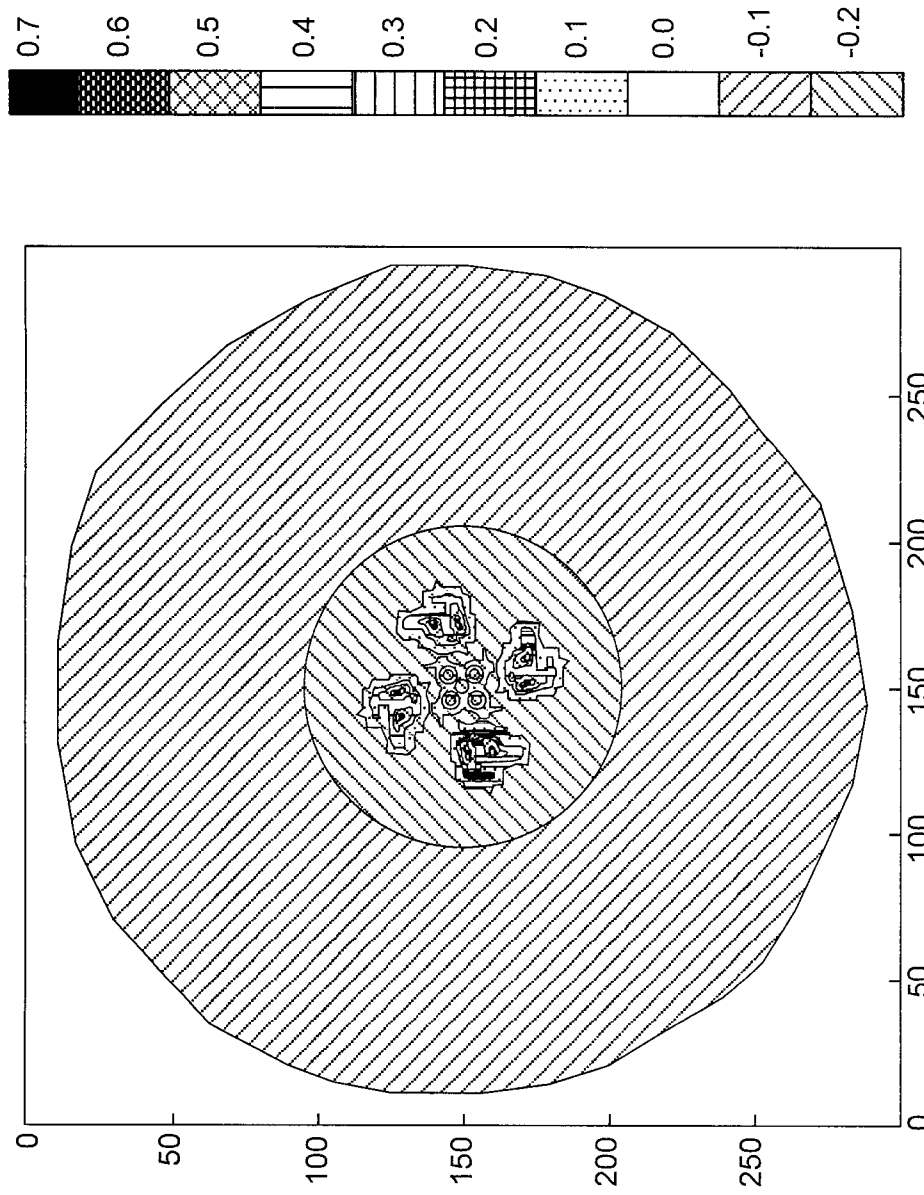


FIG. 3C

6 / 18

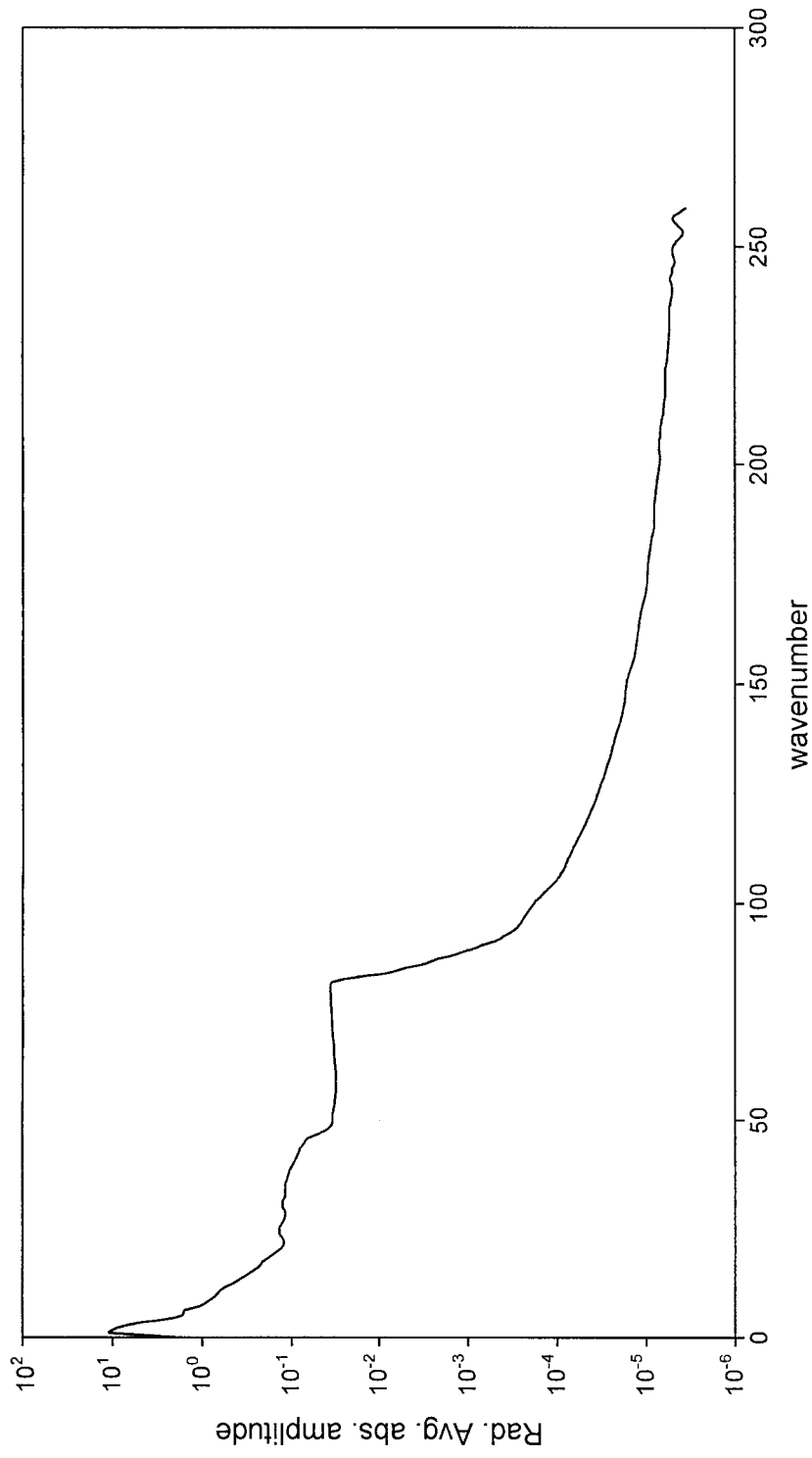


FIG. 4

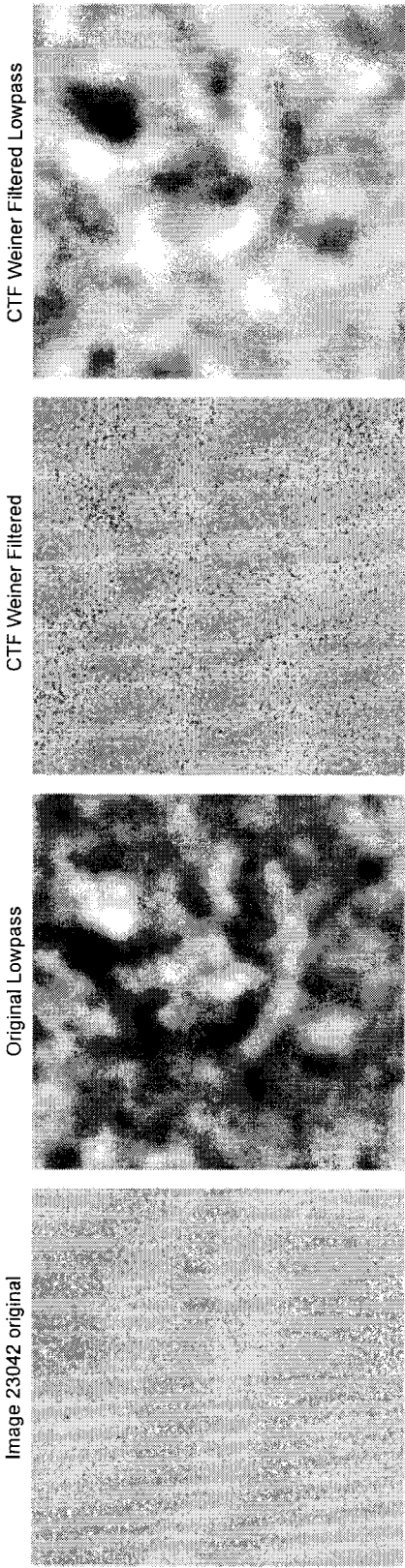


FIG. 5A

FIG. 5B

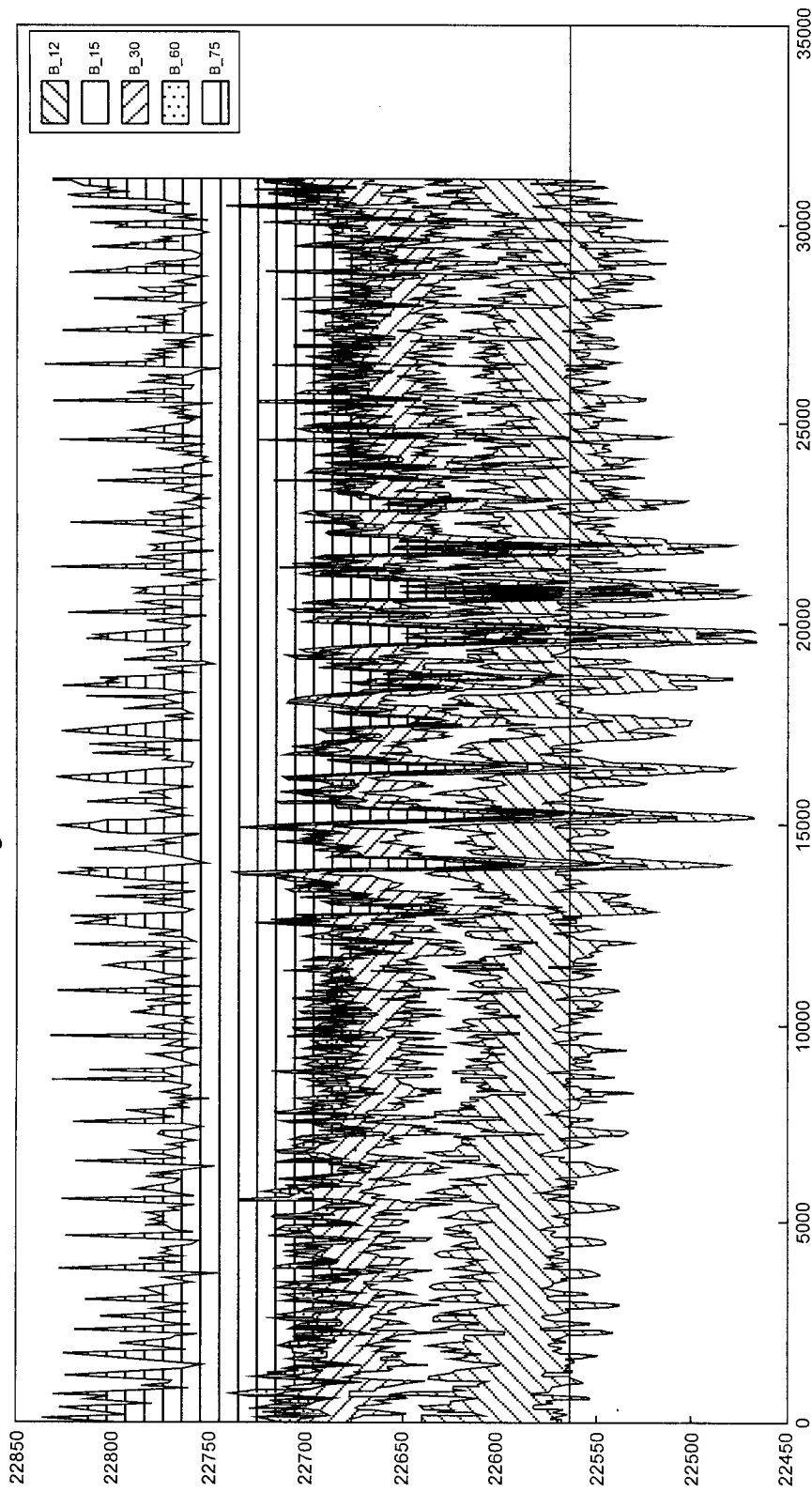
FIG. 5C

FIG. 5D

FIG. 5

8 / 18

Bounds for alignment over Pose -TRPA1 23042

**FIG. 6**

9 / 18

Bounds for alignment over Pose -TRPA1 23042

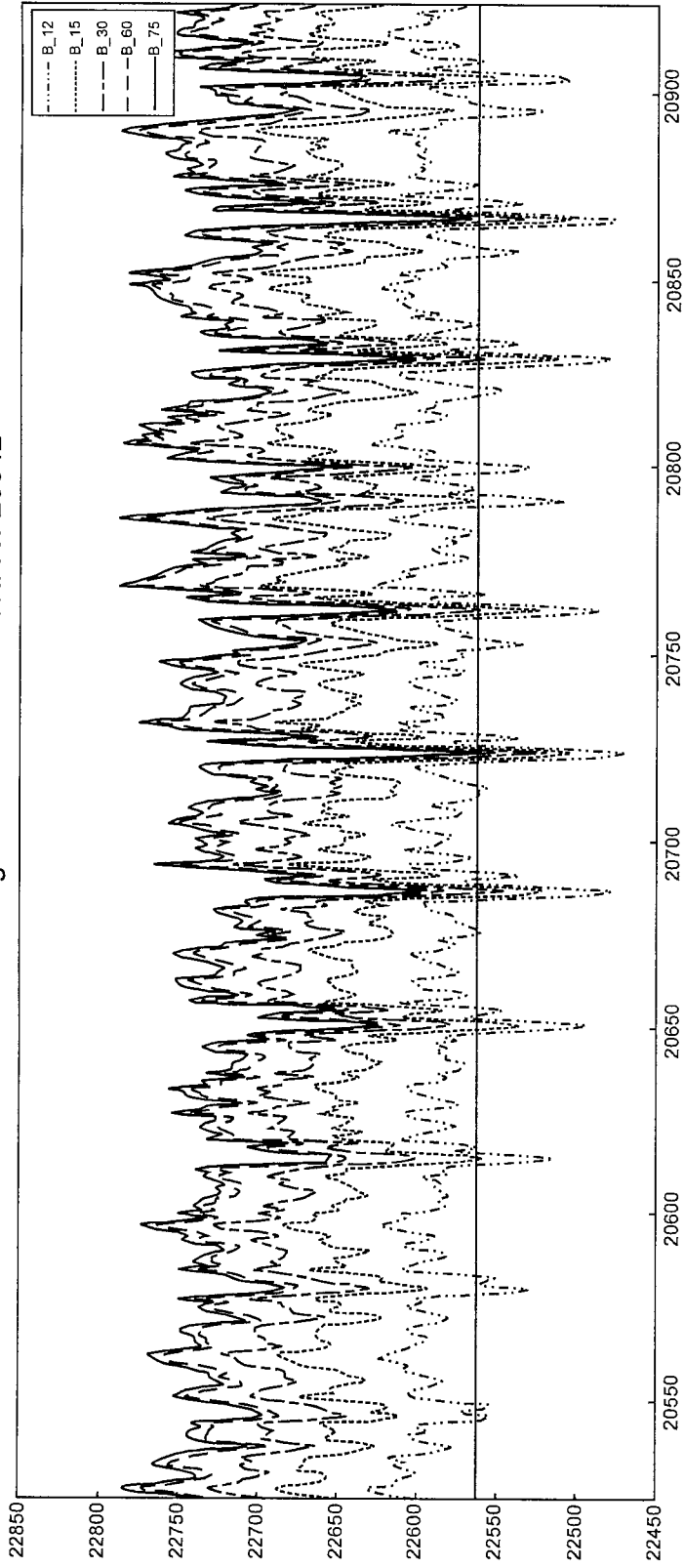
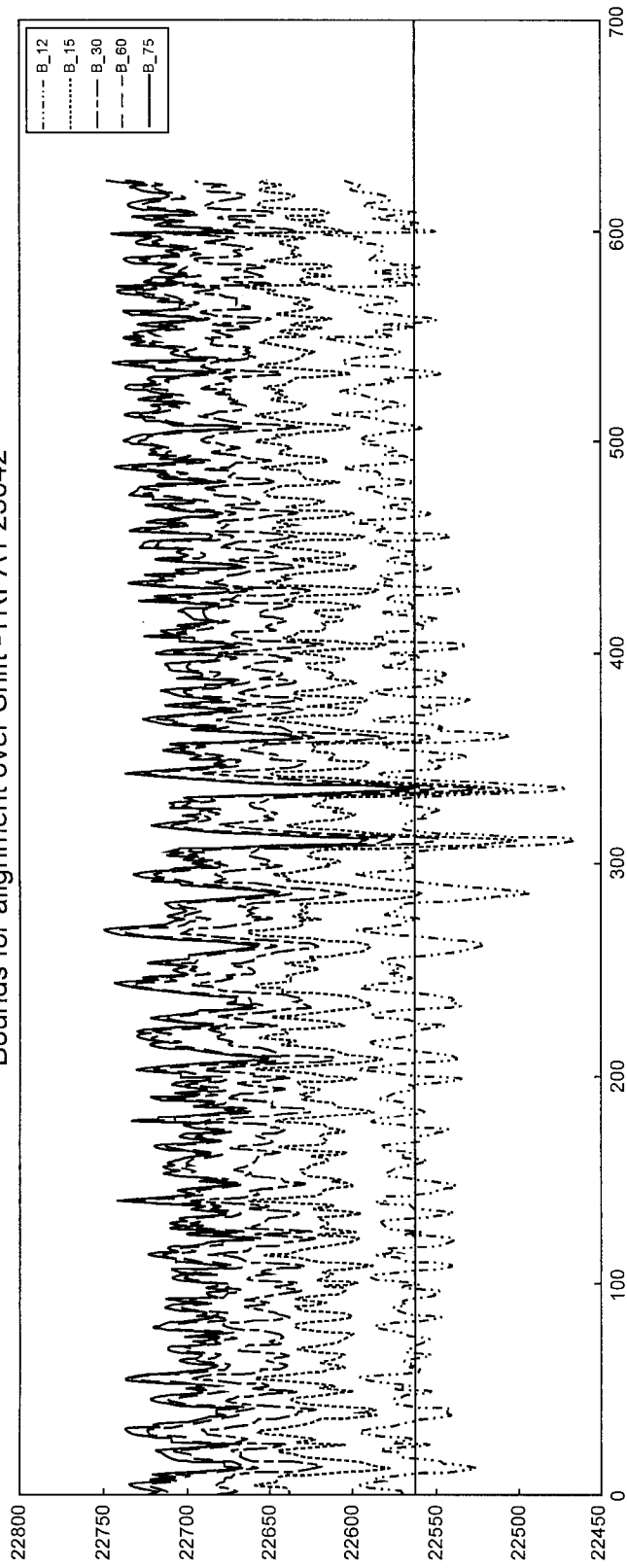


FIG. 7

10 / 18

Bounds for alignment over Shift -TRPA1 23042

**FIG. 8**

11 / 18

150

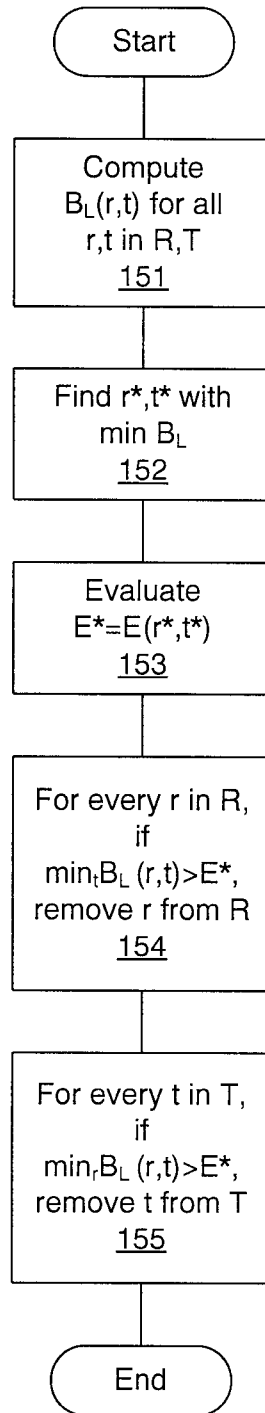


FIG. 9

12 / 18

200

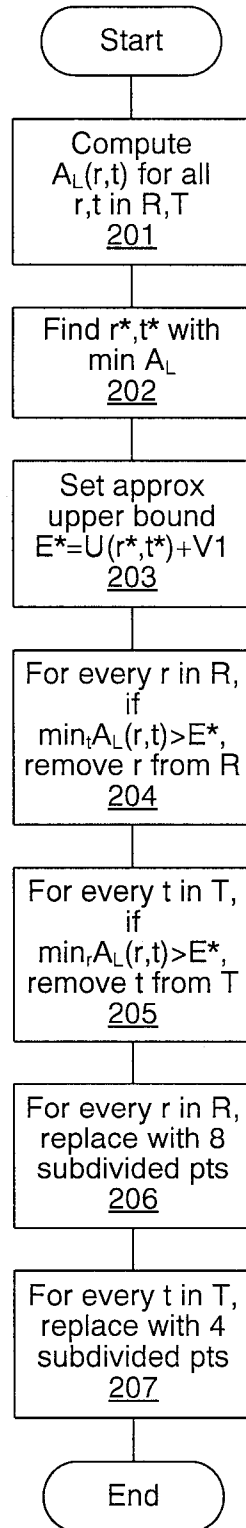
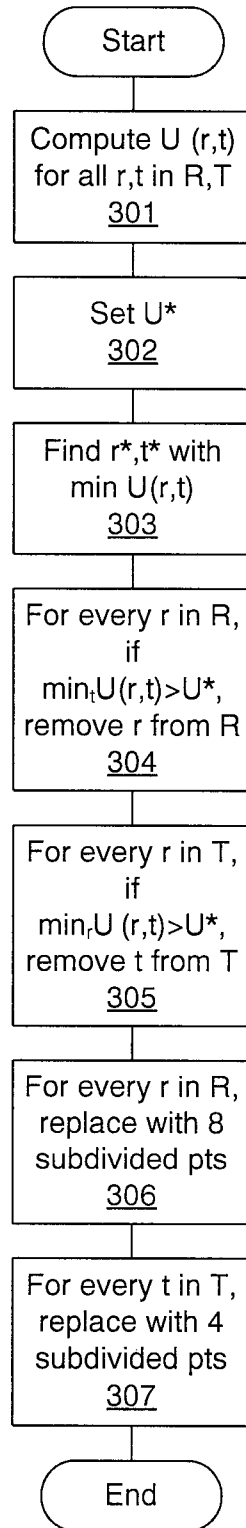


FIG. 10

13 / 18

300

**FIG. 11**

14 / 18

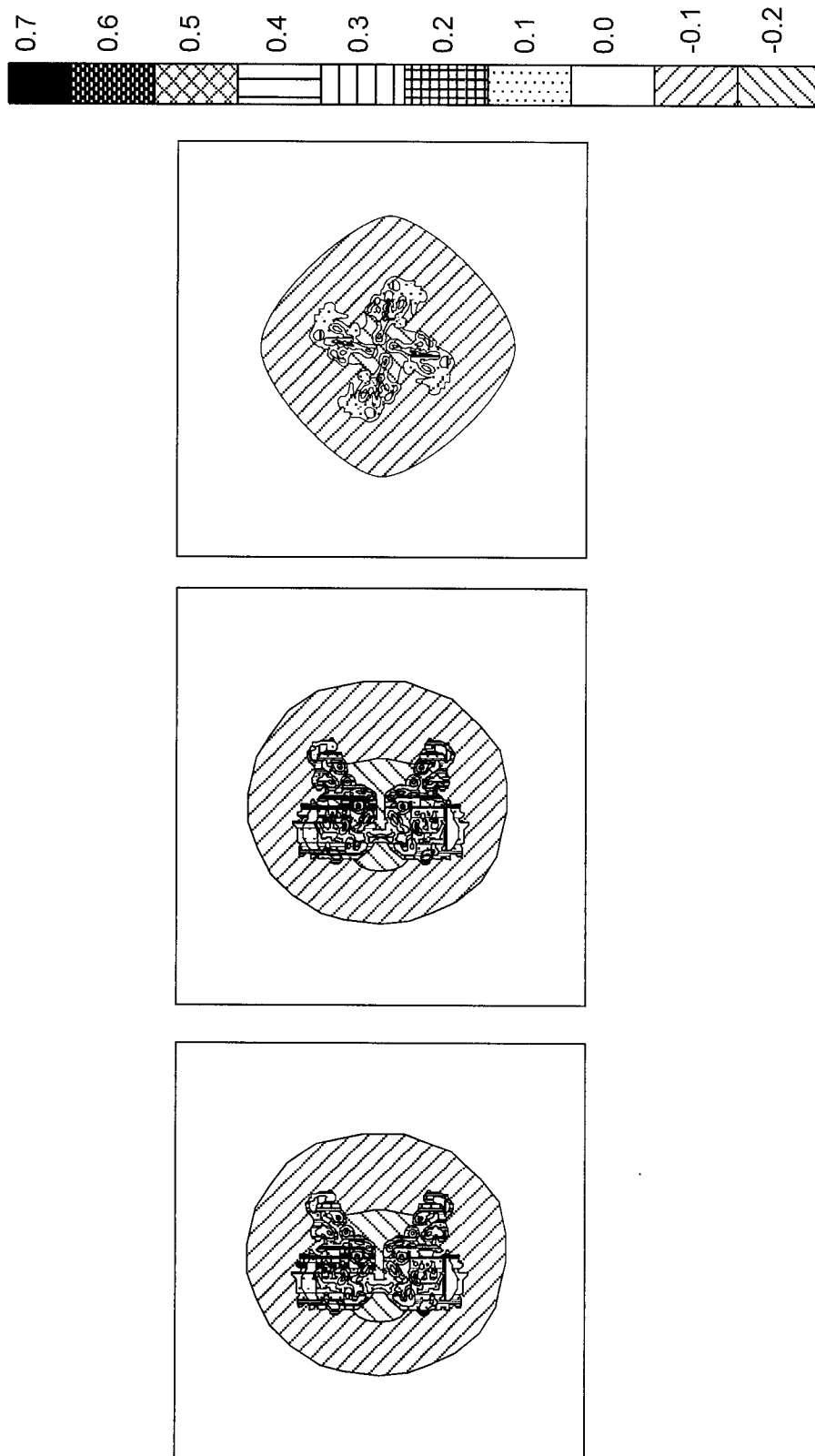


FIG. 12

15 / 18

GSFSC Resolution: 3.473A

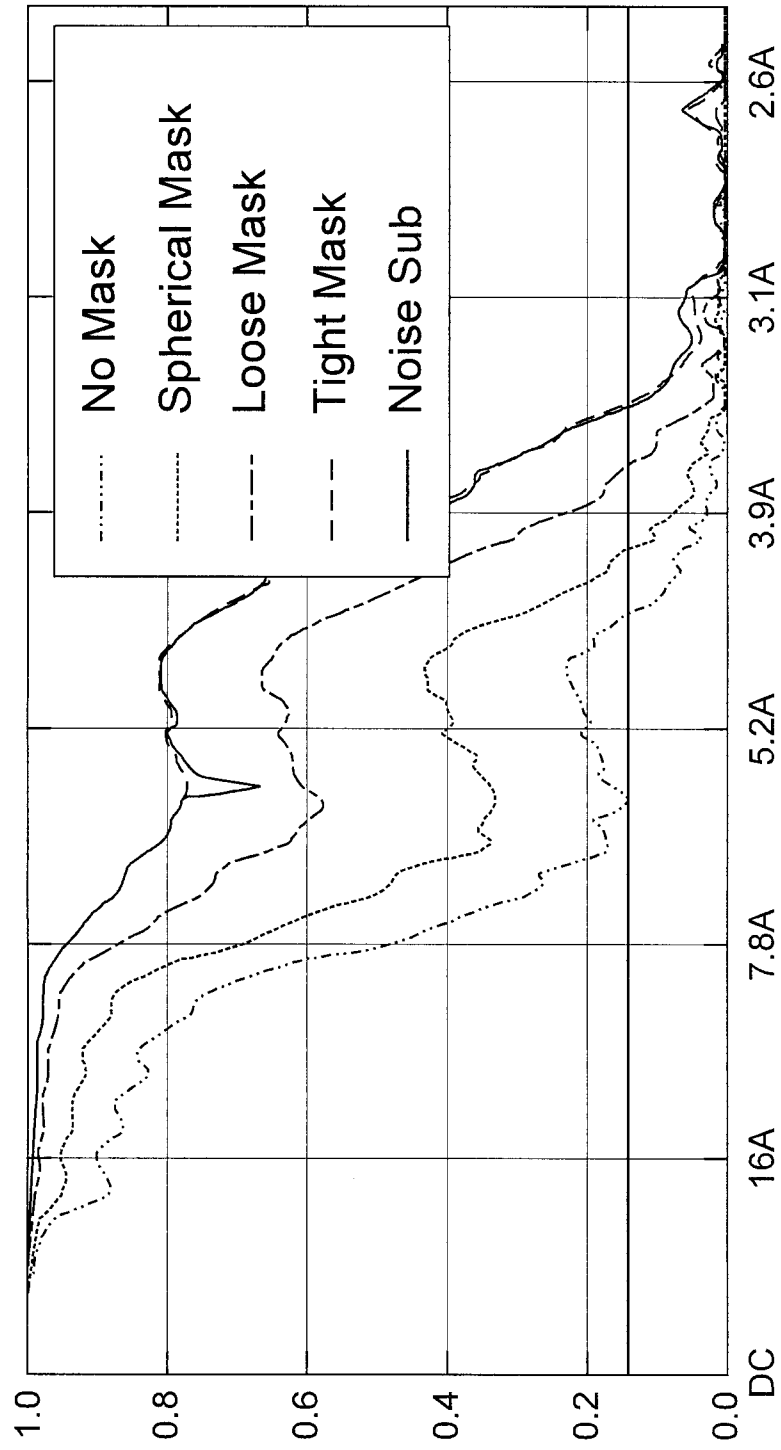


FIG. 13

16 / 18

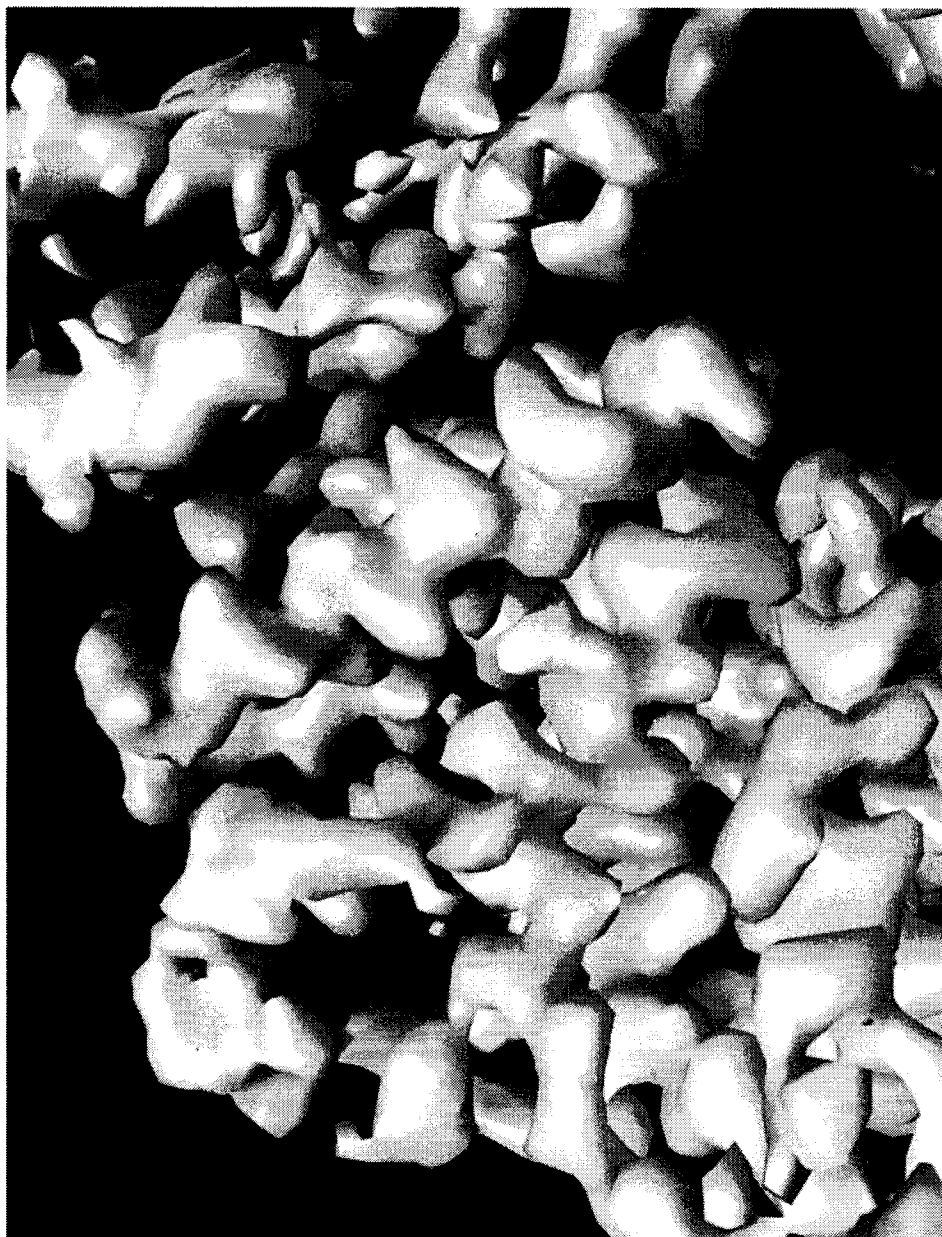


FIG. 14

17 / 18

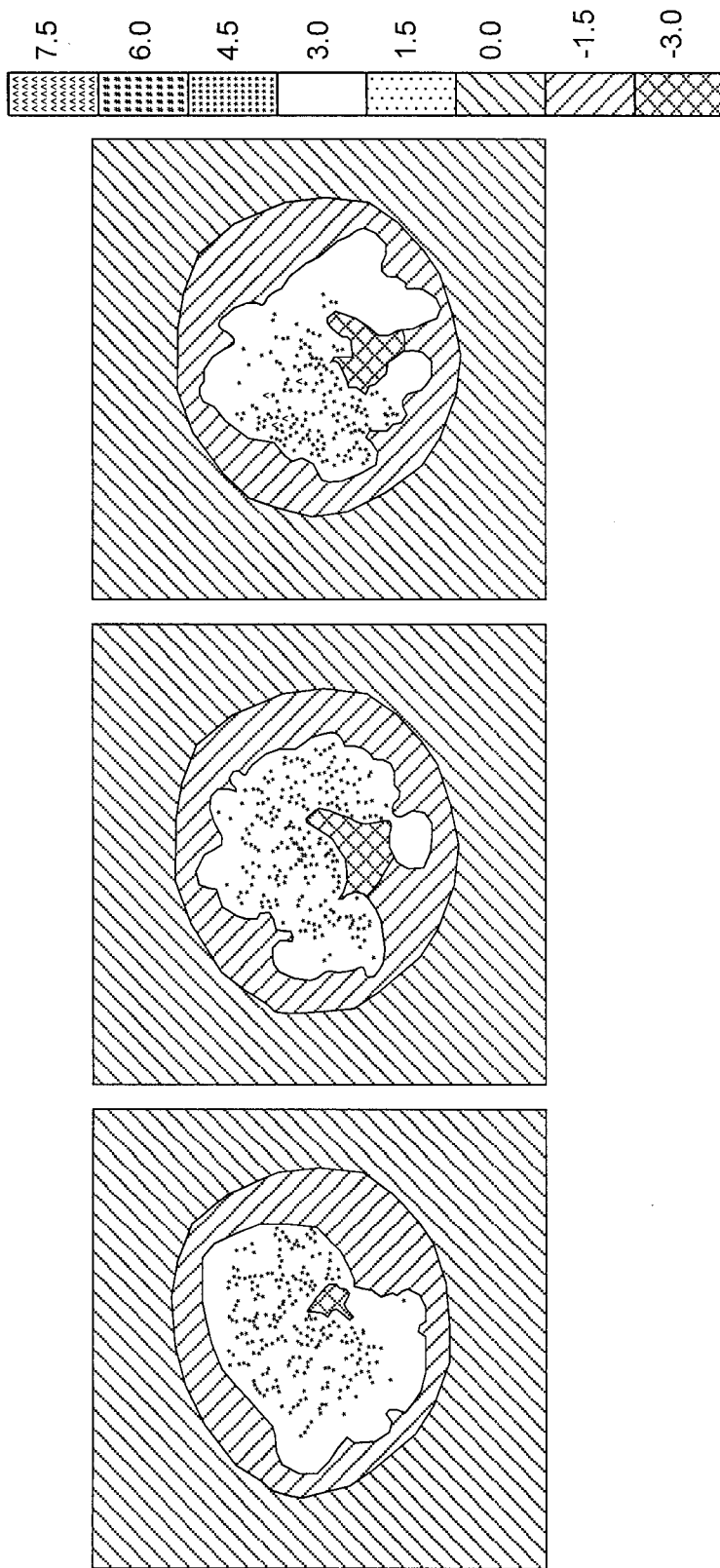


FIG. 15

18 / 18

GSFSC Resolution: 3.185A

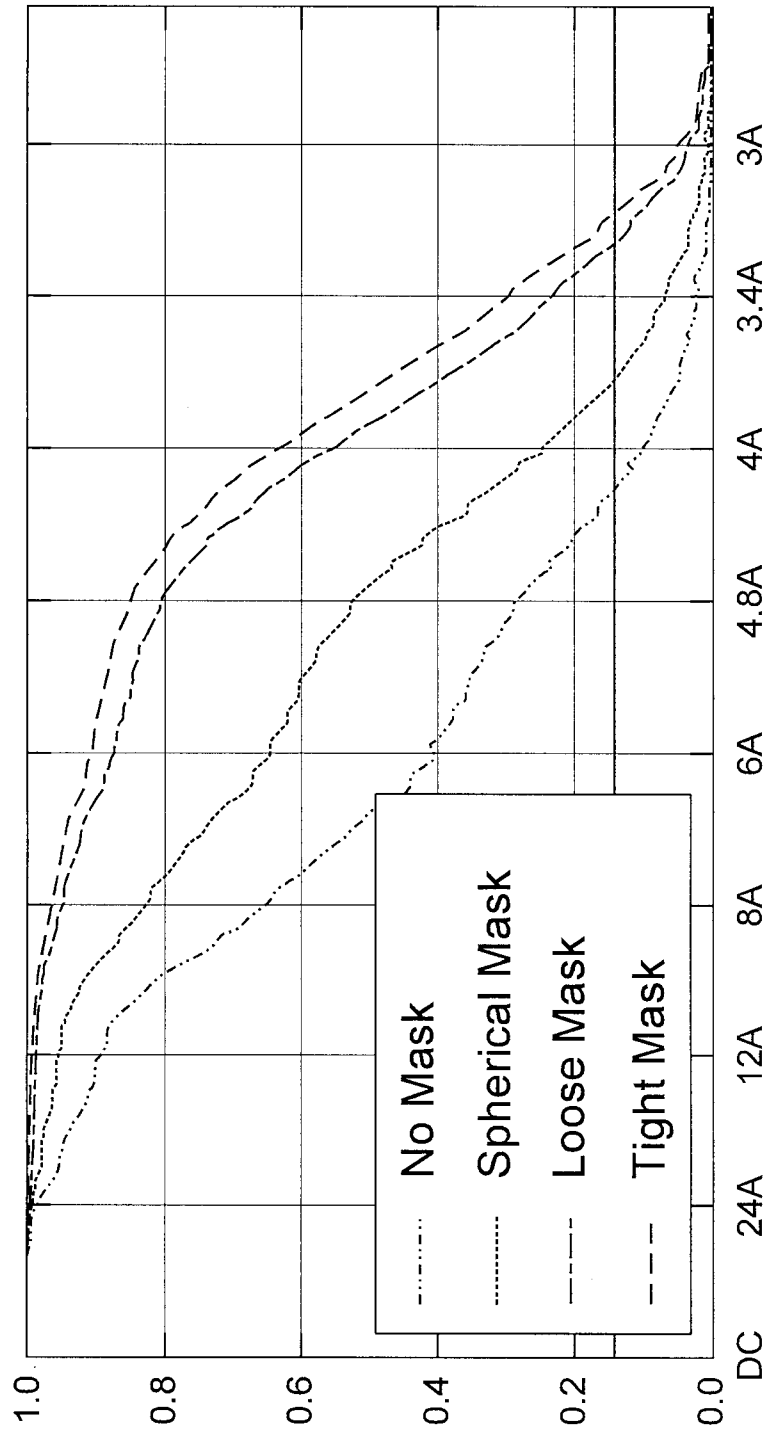


FIG. 16

