



(51) International Patent Classification:

A61B 6/00 (2006.01) G06T 5/50 (2006.01)
A61B 6/03 (2006.01) G06F 19/00 (2011.01)

(21) International Application Number:

PCT/US2012/024775

(22) International Filing Date:

11 February 2012 (11.02.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

13/025,500 11 February 2011 (11.02.2011) US

(71) Applicant (for all designated States except US): MICROSOFT CORPORATION [US/US]; One Microsoft Way, Redmond, Washington 98052-6399 (US).

(72) Inventors: KONUKOGLU, Ender; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). PATHAK, Sayan; c/o Microsoft Corporation, LCA - International Pat-

ents, One Microsoft Way, Redmond, Washington 98052-6399 (US). SIDDIQUI, Khan Mohammad; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). CRIMINISI, Antonio; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). WHITE, Steven; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). SHOTTON, Jamie Daniel Joseph; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). ROBERTSON, Duncan Paul; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR,

[Continued on next page]

(54) Title: IMAGE REGISTRATION

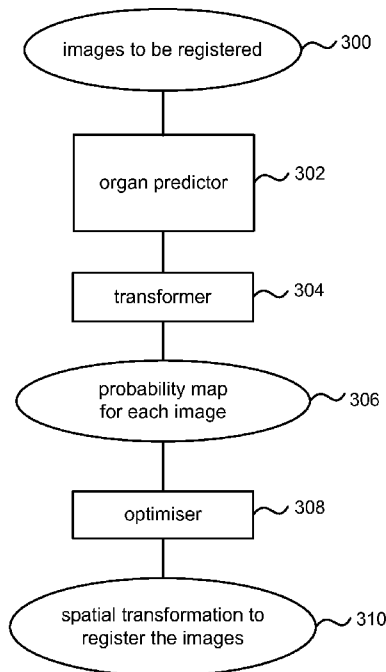


FIG. 3

(57) Abstract: Image registration is described. In an embodiment an image registration system executes automatic registration of images, for example medical images. In an example, semantic information is computed for each of the images to be registered comprising information about the types of objects in the images and the certainty of that information. In an example a mapping is found to register the images which takes into account the intensities of the image elements as well as the semantic information in a manner which is weighted by the certainty of that semantic information. For example, the semantic information is computed by estimating posterior distributions for the locations of anatomical structures by using a regression forest and transforming the posterior distributions into a probability map. In an example the mapping is found as a global point of inflection of an energy function, the energy function having a term related to the semantic information.





KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS,

SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

Published:

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))

IMAGE REGISTRATION

BACKGROUND

[0001] Image registration is used in many application domains such as medical applications, computer aided manufacturing, robotics and others. For example, two or
5 more images of the same scene may have been captured at different times and/or from different view points and/or using different image capture equipment. It may be required to put those images into registration in order to detect differences between the images which are not the result of pose variations of objects in the scene and/or of different view points of an image capture device used to capture the images.

10 [0002] Image registration in the field of medical applications is particularly useful. For example, in clinical scenarios such as patient follow-up and surgery or therapy planning. To enable a medical professional to compare medical images in an accurate manner the images need to be registered in order to remove pose variations of the patient. For example, the registration may result in alignment of corresponding anatomical
15 structures in two medical images.

[0003] There is an ongoing need to improve the accuracy and robustness of image registration systems as well as to reduce the time taken for image registration and the computational resources required.

[0004] The embodiments described below are not limited to implementations
20 which solve any or all of the disadvantages of known image registration systems.

SUMMARY

[0005] The following presents a simplified summary of the disclosure in order to provide a basic understanding to the reader. This summary is not an extensive overview of the disclosure and it does not identify key/critical elements of the invention or delineate
25 the scope of the invention. Its sole purpose is to present some concepts disclosed herein in a simplified form as a prelude to the more detailed description that is presented later.

Image registration is described. In an embodiment an image registration system executes automatic registration of images, for example medical images. In an example, semantic information is computed for each of the images to be registered comprising
30 information about the types of objects in the images and the certainty of that information. In an example a mapping is found to register the images which takes into account the intensities of the image elements as well as the semantic information in a manner which is weighted by the certainty of that semantic information. For example, the semantic

information is computed by estimating posterior distributions for the locations of anatomical structures by using a regression forest and transforming the posterior distributions into a probability map. In an example the mapping is found as a global point of inflection of an energy function, the energy function having a term related to the semantic information.

[0006] Many of the attendant features will be more readily appreciated as the same becomes better understood by reference to the following detailed description considered in connection with the accompanying drawings.

DESCRIPTION OF THE DRAWINGS

[0007] The present description will be better understood from the following detailed description read in light of the accompanying drawings, wherein:

FIG. 1 is a schematic diagram of a medical image registration system;

FIG. 2 is a schematic diagram of a pair of medical images to be registered and using either of two types of energy function which are shown schematically;

FIG. 3 is a flow diagram of an example method of image registration at a medical image registration system;

FIG. 4 is a schematic diagram of an example portion of a decision forest;

FIG. 5 is an example of a training image for estimating organ positions;

FIG. 6 is a flow diagram of an example method of training a regression forest;

FIG. 7 is a flow diagram of an example method for predicting organ locations using a trained regression forest;

FIG. 8 is a flow diagram of an example method of computing probability maps;

FIG. 9 is an example of probability maps computed from an input image;

FIG. 10 is a flow diagram of an example method of image registration;

FIG. 11 is an example of a registered image;

FIG. 12 illustrates an exemplary computing-based device in which embodiments of a medical image registration system may be implemented.

Like reference numerals are used to designate like parts in the accompanying drawings.

DETAILED DESCRIPTION

[0008] The detailed description provided below in connection with the appended drawings is intended as a description of the present examples and is not intended to represent the only forms in which the present example may be constructed or utilized. The description sets forth the functions of the example and the sequence of steps for

constructing and operating the example. However, the same or equivalent functions and sequences may be accomplished by different examples.

5 [0009] Although the present examples are described and illustrated herein as being implemented in a medical image registration system, the system described is provided as an example and not a limitation. As those skilled in the art will appreciate, the present examples are suitable for application in a variety of different types of image registration systems.

10 [0010] The techniques below are described with reference to medical images, which can be two or three dimensional images representing an internal structure of a (human or animal) body, or a sequence of such images. The medical images may be two dimensional or higher dimensional volumetric images representing the internal structure of a human or animal body (or a sequence of such images). Volumetric images and are formed of voxels. A voxel in a three dimensional volumetric image is analogous to a pixel in a two dimensional image and represents a unit of volume.

15 [0011] Although described with reference to medical images the techniques described below may also be applied in other fields, for example registration of satellite data, registration of sonar data, topographical modeling or other computer vision processes.

20 [0012] The term 'image element' is used herein to refer to a pixel in a two dimensional image, a voxel in a three or higher dimensional image or time varying sequence of images, or groups of pixels or voxels such as blobs, patches, or other collections of two or more pixels or voxels. Each image element has one or more values associated with it. Each value represents a property such as intensity or color. The property can depend on the type of medical image device generating the image. In an example the image intensities can be related to the density of the tissue at a given portion of the image or the proportion of water present in the material.

25 [0013] FIG. 1 is a schematic diagram of a medical image registration system comprising a computer implemented image registration system 106. In an example a user 110, for example a doctor, may wish to examine multiple images 102, 104, for example for patient follow up or to plan surgery.

30 [0014] The image registration system 106 is arranged to take a reference image 102 of two or higher dimensions and a target image 104 which is also of two or higher dimensions and perform a transformation so that the second medical image is registered to

the first medical image or vice versa. In an example the reference image and the target image may depict the same patient. The images may be taken using different equipment (e.g. be multi-modal) or at different times or from a different angle. The image registration system enables these variations to be removed before medical staff, other users
5 or automated processes compare the images.

[0015] In an example a plurality of images may be registered to the reference image. For example a doctor may wish to perform a multi-modal examination of a patient using e.g. an X-ray, a computed tomography (CT) scan and a positron emission tomography (PET) scan. Once the images have been registered to the reference image, the
10 user 110 is able to compare the registered images 108 output from the image registration system 106 in order to look for any differences. For example the user may examine the registered images to look for differences. In an example, the registered images 108 may show a change in the size or shape of an organ over time, which may indicate injury, illness or another problem with the organ. Multi-modal registered images may also be
15 used for further automated analysis. For example, scientific data mining or fast visualization. An automated organ recognition process may be applied to the registered images with multi-modal features in order to achieve higher accuracy than using such an organ recognition process prior to image registration.

[0016] In some examples described below the images are of at least three
20 dimensions although this is not essential. The image registration system 106 may also be used with two-dimensional images.

[0017] The image registration system may display the output images in a plurality of different ways. In an example the registered images 108 may display the reference image 102 as a black and white or color image, with the target image 104 overlaid as a
25 contour map or a partially transparent image. In another example the registered images may be displayed as an image sequence. The registered images 108 may be displayed in a pre-specified manner or the user 110 may select different ways of displaying the image. For example the user 110 may choose to first display a target image as a contour overlay on the reference image and then choose to display two or more of the images in sequence.
30 For example blinking the images. In some embodiments the estimated uncertainty in the image registration may also be displayed by varying the color or degree of transparency for example. This uncertainty value may be provided as a variance or standard deviation

value for example and provides an indication of how certain the image registration system 106 is that the registration of the target image 104 to the reference image 102 is correct.

[0018] In an example the medical images may be from a CT scanner 100.

However, the images may be produced by many types of medical equipment such as
5 magnetic resonance imaging (MRI) scanners, computed tomography (CT) scanners, single
photon emission computed tomography (SPECT) scanners, positron emission tomography
(PET) scanners and ultrasound scanners.

[0019] FIG. 2 is a schematic diagram of a pair of medical images to be registered
and two types of energy function which are shown schematically. Many common
10 techniques for registration of images cast the linear registration problem as the
minimization of a global energy function based on intrinsic image properties, for example
an energy function based on the image intensity $E(I)$. In an example of this technique a
target image 202 can be registered to a reference image 200 using a minimization of the
energy function 204. However, the minimization step of algorithms which minimize an
15 energy function based on only image intensities and without using semantic information
about the image content can become caught in local minima, leading to unsatisfactory
results.

[0020] In another example the images 200, 202 can be registered using an
optimization of an energy function 206 based on an intrinsic image property and other
20 semantic information $E(I + \text{semantic information})$. For example, the semantic
information may be information about the types of anatomical structures or organs in a
medical image and certainty of that information. The semantic information may be a
probability measure, for example the probability of an organ being located in a certain part
of an image. In other examples the semantic information may be classification
25 information, segmentation information or other information about the content of the
image. The addition of semantic information as a term in the energy function enables an
energy function to be computed which has a more clearly defined global point of
inflection (minimum or maximum) to be found and improves the robustness of image
registration.

30 [0021] In an example an automatic image registration system comprises an input
arranged to receive a first image and a second image both images being of objects and
with at least part of one object being common to both images. For example, the images
may be medical images of a patient's thorax both showing the patient's lungs but from

slightly different viewpoints. The system may have a processor arranged to compute semantic information for each of the first and second images, the semantic information comprising information about the types of objects in the images and the certainty of that information. In some cases this semantic information about the image content is provided
5 as probability maps described in more detail below. However, other formats may be used for the semantic information. The processor may be further arranged to find a mapping to register the first and second images which takes into account the intensities of the image elements of the images and also takes into account the semantic information in a manner which is weighted by the certainty of that semantic information. This gives improved
10 quality of image registration.

[0022] In an example, a method of automatic image registration comprises receiving a first image and a second image both images being of objects and with at least part of one object being common to both images. For each of the first and second images a probability map is computed comprising, for each image element, a probability that the
15 image element is of a specified object. A mapping may then be found to register the first and second images by optimizing an energy function which is a function of the intensities of the first and second images and also of the probability maps.

[0023] FIG. 3 is a flow diagram of an example method of image registration at the medical image registration system. For example the image registration system of FIG. 1.
20 The images to be registered 300 are input to an organ predictor 302. The organ predictor may predict the location of an organ or other anatomical structure in an image. For example the organ predictor 302 may be arranged to provide an estimated region or bounding box for each organ it detects, together with associated uncertainties. In an example where the input images are three-dimensional images the estimated region is a
25 volume which tightly contains or encompasses an organ; it may be a box or any other 3D form. In other examples the organ predictor may be configured to label or segment the images.

[0024] In an embodiment the organ predictor comprises a regression forest. A regression forest is an ensemble classifier comprising a plurality of trained regression
30 trees. An ensemble of many randomly trained regression trees (a regression forest) yields improved generalization over a single tree which may suffer from over-fitting. Regression trees differ from classification trees in that they provide real valued continuous output as opposed to a class to which an object belongs. During the training process, the number of

trees is fixed. In one example, the number of trees is three, although other values can also be used. The choice of the number of trees and the depth of those trees depends on the application domain as well as factors such as the computational resources available.

[0025] The use of a regression forest to predict organ locations is described in more detail with reference to FIG's 4-7. In a non-exhaustive list of examples the organ predictor may alternatively be; an alternative type of trained decision forest, for example a classification forest system; a vector machine such as a support vector machine (SVM); a relevance vector machine (RVM); in addition adaptive boosting techniques (e.g. Ada boost) may be used.

10 [0026] Predictions received from the organ predictor 302 for the location of an organ in an image 300 are used as input to a transformer 304. The transformer 304 is used to place the semantic information into a form which may be readily used as part of a process for finding a mapping between the images. For example, the transformer 304 may transform the output of the organ predictor into a probability map 306 for each image. A
15 probability map can be considered as an image (which may be of 2, 3 or higher dimensions) where each image element location holds one or more values, each value corresponding to the probability that that image element depicts a specified organ. Each element of the probability map may correspond to an image element in the input image, or may be of different resolution. For example each element in the probability map may be an
20 average of the probability at several image elements. In some examples herein the probability map is a voxel-wise probability map. A color table of a probability map may be used to visualize the probability value. An example of probability maps is shown in FIG. 9.

[0027] In an example, the probability maps 306 for each image are used by an
25 optimizer 308 to compute a spatial transformation 310 to register the input images. For example the optimizer 308 attempts to find an optimal solution of an energy function as is described above with reference to FIG. 2. This energy optimization problem may be formulated from the image intensity and the semantic information in order to find a global point of inflection. In an example the registration problem is an energy optimization
30 problem. An example of energy optimization is described with reference to FIG. 10.

[0028] FIG. 4 is a schematic diagram of an example portion of a random forest which may be a random regression forest, random decision forest or other type of random forest. The illustrative decision forest of FIG. 4 comprises three decision trees: a first tree

400 (denoted tree Ψ_1); a second tree 402 (denoted tree Ψ_2); and a third tree 404 (denoted tree Ψ_3). Each decision tree comprises a root node (e.g. root node 406 of the first decision tree 400), a plurality of internal nodes, called split nodes (e.g. split node 408 of the first decision tree 400), and a plurality of leaf nodes (e.g. leaf node 410 of the first decision tree 400).

5 [0029] Highly non-linear mappings are handled by splitting the original problem into a set of smaller problems which can be addressed with simple predictors. FIG. 4 shows an illustrative 1D example where the goal is to learn an analytical function to predict real-valued output. A single linear function fits the data badly. For example a single linear function 412 as output by root node 406 does not fit the data. A set of training pairs (x, y) 410 is given to a tree, for example tree 400, at the root node, for example root node 406. Each node in the tree is designed to split the data so as to form clusters where accurate prediction can be performed with simpler models (e.g. linear in this example). Using more tree levels may yield more accurate fit of the regressed model. For example 15 the training pairs 410 may be modeled 414 using a combination of linear regressors from nodes 416 to 422 which fit the data well. In operation, each root and split node of each tree performs a binary test on the input data and based on the result directs the data to the left or right child node. Formally each node performs the test $\xi > f(x) > \tau$, with ξ, τ scalars. Based on this result each data point is sent to the left or right child node 408.

20 [0030] In an example where the decision forest is a regression forest the leaf nodes store continuous parameters η_s characterizing each regressor. The regressor may be linear, constant, Gaussian, quadratic or any other functional form. This is in contrast to classification or decision forests where the predicted variables are discrete. Complex non-linear mappings may be modeled via a hierarchical combination of many simple linear 25 regressors. More tree levels yield smaller clusters and smaller fit errors but may overfit the input data. The manner in which the parameters used by each of the split nodes are chosen and how the leaf node probabilities are computed is described with reference to FIG's 5 and 6.

[0031] The tree based regressor allows multiple anatomical structures to be dealt 30 with simultaneously which encourages feature sharing between anatomical structures and therefore better generalization. For example, the presence of a lung may be indicative of the presence of the heart. In contrast to classification based approaches which can localize the organ centers but not the organ extent, the continuous model estimates the position of

the walls of the bounding box containing each organ, thus achieving simultaneous organ localization and extent estimation.

[0032] As described in the examples below a regression forest algorithm can incorporate atlas information within the compact tree based model. Atlas-based techniques have conceptual simplicity. An atlas is a hand-classified image, which is mapped to a subject image by deforming the atlas until it closely resembles the subject. This technique is therefore dependent on the availability of good atlases. Robustness can be improved by multi-atlas based techniques, however this can be computationally inefficient. In addition, the conceptual simplicity of such algorithms is in contrast to robust cross-image registration. By incorporating atlas information into a regression model greater efficiency and robustness is achieved than using multiple atlases and anatomy localization is achieved more quickly.

[0033] In the example described with reference to FIG.4 a regression forest was used to estimate a two-dimensional continuous vector representing a line. However, a regression forest may be used to estimate variables with a plurality of dimensions. In an example where the locations of organs in a volumetric image are being estimated the output is a six dimensional vector \mathbf{b}_c per organ, for a total of $6|C|$ continuous parameters. In an embodiment the anatomical structures which can be identified include but are not limited to $C = \{\text{heart, liver, spleen, left lung, right lung, left kidney, right kidney, gall bladder, left pelvis, right pelvis}\}$.

[0034] FIG. 5 is an example of a training image for estimating organ positions. Note that the schematic diagram of FIG. 5 is shown in two dimensions only, for clarity, whereas an example volumetric image is three-dimensional. The medical image 500 comprises a representation of several organs, including a kidney 502, liver 504 and spinal column 506, but these are only examples used for the purposes of illustration. Other typical organs that can be shown in images and identified using the technique described herein include (but are not limited to) the head, heart, eyes, lungs, and major blood vessels. A bounding box 508 is shown drawn (in dashed lines) around the kidney 502. Note that in the illustration of FIG. 5 the bounding box 508 is only shown in two dimensions, whereas in a volumetric image the bounding box 508 surrounds the kidney 502 in three dimensions.

[0035] Each voxel v (for example v_1, v_2 510, 512) in the training volume can be associated with an offset $\mathbf{d}_c(\mathbf{v}) = (d_c^l, d_c^r, d_c^h, d_c^f, d_c^a, d_c^p) \in \mathbb{R}^6$ with respect to the bounding box \mathbf{b}_c 508 for each class $c \in \mathcal{C}$. Examples of offsets are shown in FIG. 5 for two voxels v_1 510, and v_2 512. The offset for v_1 510 is given by

5 $\mathbf{d}(\mathbf{v}_1) = (d_1^l, d_1^r, d_1^h, d_1^f)$ and for v_2 512 is given by $\mathbf{d}(\mathbf{v}_2) = (d_2^l, d_2^r, d_2^h, d_2^f)$ the other two components of the offset vectors (d_1^a, d_1^p) and (d_2^a, d_2^p) which would be present in a 3-D volumetric image are not shown in the two dimensional representation in FIG. 5.

[0036] In an embodiment all voxels in a test CT volume contribute with varying confidence to estimating the position of the walls of a bounding box. In other
10 embodiments only a selection of voxels may contribute. Some clusters of voxels may predict the position of an organ with high confidence. This may be the case even if the voxel does not comprise part of the organ, for example, voxel clusters that are part of the ribs or vertebrae may predict the position of the heart with high confidence. During testing
15 these clusters may be used as references for the location of anatomical structures. Voxels can be clustered together based on their appearance, their spatial context, their intensity value, their confidence levels or any other appropriate measure. The feature selection and parameter regression can be undertaken using a multi-class random regression forest e.g. an ensemble of regression trees trained to predict the location and size of all desired
20 anatomical structures simultaneously.

[0037] FIG. 6 is a flow diagram of an example method of training a regression forest. FIG. 6 describes one example method of training a regression forest. Other methods may be used such as breadth first training where a whole level of the tree is trained at once or other methods. A training set of labeled volumes can be created with
25 associated ground-truth organ bounding boxes. In an example the training set is created from a subset of all the volumes. The training set may be created by hand drawing a bounding box 600 which may be a cuboid in the case of a 3D image or a rectangle in the case of a 2D image, centered on the organ of interest. In an example there may be more than one organ of interest per image. The bounding box can also be extended in the
30 temporal direction in the case of a sequence of images.

[0038] The bounding box may be drawn using a dedicated annotation tool, which is a software program enabling fast drawing of the bounding boxes from different views of

the image (e.g. axial, coronal, sagittal and 3D views). The bounding boxes may be drawn manually. In some embodiments radiologists can be used to validate that the labeling is anatomically correct.

[0039] The number of regression trees to be used in the regression forest is set 602.

5 A regression forest is a collection of regression trees. The forest is composed of T trees denoted $\Psi_1, \dots, \Psi_t, \dots, \Psi_T$ with t indexing each tree. During the training process the number of trees is fixed. In an example the number of trees is three (for example the regression forest of FIG. 4) but other values may also be used. All trees are trained in parallel. However, the trees may also be trained separately.

10 [0040] At each tree in the forest the root node (e.g root node 406) is selected 604. All image elements from each of the training images are then selected 606. An image element in an image V is defined by its co-ordinates $\mathbf{v} = (x, y, z)$.

[0041] The manner in which the parameters used by each of the split nodes are chosen and how the continuous parameters are stored and computed is now described.

15 [0042] A random set of test parameters are generated 608 for use by the binary test performed at the root node 406. In one example the binary test is of the form:

$\xi > f(\mathbf{v}; \theta) > \tau$, such that $f(\mathbf{v}; \theta)$ is a function applied to image element at position \mathbf{v} and its surrounding image elements with parameters θ and with the output of the function compared to threshold values ξ and τ . If the result of $f(\mathbf{v}; \theta)$ is in the range between ξ

20 and τ then the result of the binary test is true, otherwise the result of the binary test is false. In other examples, only one of the threshold values can be used such that the binary test is true if the result of $f(\mathbf{v}; \theta)$ is greater than (or less than) a threshold value. In the example described here, the parameter θ defines a visual feature of the image.

[0043] Anatomical structures can be difficult to identify in medical images

25 because different organs can share similar intensity values, e.g. similar tissue density in the case of CT and X-Ray scans. Thus, local intensity information is not sufficiently discriminative to identify organs and as described above this may lead to the optimization of the energy function becoming stuck in local minima.

[0044] The parameters θ for the function $f(\mathbf{v}; \theta)$ are randomly generated during

30 training. The process for generating the parameters θ comprises generating, for example, a randomly-sized box (a cuboid box for 3D images, or a rectangle for 2D images, both of which can be extended in the time-dimension in the case of a sequence of images) and a

spatial offset value. All dimensions of the box are randomly generated. The spatial offset value is in the form of a two- or three-dimensional displacement. In other examples, the parameters θ can further comprise one or more additional randomly generated boxes and spatial offset values. In alternative examples, differently shaped regions (other than boxes) or offset points can be used. In an example, only a single signal channel (e.g. intensity only) is used for all boxes. In other examples the channel may be the magnitude of the intensity gradient or a more complex filter can be used.

[0045] Given the above parameters θ , the result of the function $f(\mathbf{v};\theta)$ is computed by aligning the randomly generated box with the image element of interest \mathbf{v} such that the box is displaced from the image element \mathbf{v} in the image by the spatial offset value. The value for $f(\mathbf{v};\theta)$ is then found by summing the values for the signal channel for the image elements encompassed by the displaced box (e.g. summing the intensity values for the image elements in the box). Therefore, for the case of a single box,

$f(\mathbf{v};\theta) = \sum_{\mathbf{q} \in F} I(\mathbf{q})$, where \mathbf{q} is an image element within box F . This summation is

normalized by the number of pixels in the box. In the case of two boxes, $f(\mathbf{v};\theta)$ is given by: $f(\mathbf{v};\theta) = |F_1|^{-1} \sum_{\mathbf{q} \in F_1} I(\mathbf{q}) - |F_2|^{-1} \sum_{\mathbf{q} \in F_2} I(\mathbf{q})$, where F_1 is the first box, F_2 may be a second box or, in other examples, F_2 may be an empty set for unary features. Again, these two summations are normalized separately by the respective number of image elements in each box.

[0046] The result of the binary test performed at the root node or split node determines which child node an image element is passed to. For example if the result of the binary test is true, the image element is passed to a first child node, whereas if the result is false, the image element is passed to a second child node.

[0047] The random set of test parameters generated comprise a plurality of random values for the function parameters θ and the threshold parameter values ξ and τ . In order to inject randomness into the decision trees, the function parameters θ of each split node are optimized over only a randomly sampled subset Θ of all possible parameters. For example, the size of the subset can be 100. This is an effective way of increasing generalization.

[0048] Every combination of test parameter 610 is applied to each image element in the training images. In other words, all available values $\theta \in \Theta$ are tried one after

another, in combination with all available values of ξ and τ for each image element in each training image. For each combination, the information gain, also known as the relative entropy is calculated by selecting parameters maximizing information 612 given.

[0049] For example, as described above, each voxel v in the training volume is

5 associated with an offset $\mathbf{d}_c(\mathbf{v}) = (d_c^l, d_c^r, d_c^h, d_c^f, d_c^a, d_c^p) \in \mathbb{R}^6$ with respect to the bounding box \mathbf{b}_c for each class $c \in \mathbf{C}$ where $\mathbf{b}_c(\mathbf{v}) = \hat{\mathbf{v}} - \mathbf{d}_c(\mathbf{v})$ and

$\hat{\mathbf{v}} = (v_x, v_x, v_y, v_y, v_z, v_z)$. Node optimization may be executed by optimizing an

information gain measure. An example of an information gain measure is

$IG = H(\mathbf{S}) - \sum_{i \in \{L,R\}} \omega_i H(\mathbf{S}_i)$ where H denotes entropy, \mathbf{S} is the set of training

10 points reaching the node and L,R denote the left and right child nodes.

[0050] In an example where the decision forest is a classification forest the entropy is measured over discrete class labels. In contrast, in an example where the decision forest is a regression forest the purity of the probability density of the real valued predictions is measured. For a single class c the distribution of the vector \mathbf{d}_c can be modeled at each

15 node as a multivariate Gaussian. For example $p(\mathbf{d}_c) = \mathbf{N}(\mathbf{d}_c; \bar{\mathbf{d}}_c \Lambda_c)$, with the matrix Λ_c encoding the covariance of \mathbf{d}_c for all points in \mathbf{S} . In other examples the regression forest may alternatively be used with a non-parametric distribution.

[0051] The differential entropy of a multivariate Gaussian can be shown to be

$H(\mathbf{S}) = \frac{n}{2}(1 + \log(2\pi)) + \frac{1}{2} \log |\Lambda_c(\mathbf{S})|$ where n is the number of dimensions. In examples

20 herein where the image are volumetric images $n = 6$, in examples where two dimensional images are described $n = 4$. However, n may take any appropriate value. The regression information gain in this example is therefore $IG = \log |\Lambda_c(\mathbf{S})| - \sum_{i \in \{L,R\}} \omega_i \log |\Lambda_c(\mathbf{S}_i)|$.

[0052] In examples where the information gain is adapted to handle multiple organs simultaneously the information gain can be adapted to

25 $IG = \sum_{c \in \mathbf{C}} (\log |\Lambda_c(\mathbf{S})| - \sum_{i \in \{L,R\}} \omega_i \log |\Lambda_c(\mathbf{S}_i)|)$. This can be rewritten as;

$\log |\Gamma_c(\mathbf{S})| - \sum_{i \in \{L,R\}} \omega_i \log |\Gamma(\mathbf{S}_i)|$ with $\Gamma = \text{diag}(\Lambda_1, \dots, \Lambda_c, \dots, \Lambda_{|C|})$ (1)

[0053] Maximizing (1) minimizes the determinant of a $6|C| \times 6|C|$ covariant matrix Γ thus decreasing the uncertainty in the probabilistic vote cast by each cluster of

voxels on each organ pose. As an alternative to information gain other criteria can be used, such as Gini entropy or 'two-ing' criterion.

[0054] In some embodiments, if the value for the maximized information is less than a fixed threshold 614 then this indicates that the further expansion of the tree will not provide significant benefit. This gives rise to asymmetrical trees which naturally stop growing when no further nodes are needed. In other examples the node stops growing when a maximum tree depth is reached or too few points reach the node. In an example the maximum depth of the trees is 7 levels, although other values can be used. In an example the maximum depth of the trees is chosen to minimize the mean error. Where further nodes will not provide significant benefit the node is set as a leaf node 616. At each leaf node the predictions for the continuous localization parameters are stored 618. For example learned mean $\bar{\mathbf{d}}$ (with $\mathbf{d} = (\mathbf{d}_1, \dots, \mathbf{d}_c, \dots, \mathbf{d}_{|C|})$) and covariance Γ .

[0055] If the value for the maximized information gain is greater than or equal to the threshold and the tree depth is less than the maximum value then the current node is set as a split node 620. If the current node is a split node it has child nodes and the process moves to training these child nodes. Each child node is trained using a subset of the training image elements at the current node. The subset of image elements sent to a child node is determined using the parameters θ^* , ξ^* and τ^* that maximized the information gain. These parameters are used in the binary test and the binary test performed 622 on all image elements at the current node. The image elements that pass the binary test form a first subset sent to a first child node and the image elements that fail the binary test form a second subset sent to a second child node. After training the j^{th} split node remains associated with the feature θ_j and the thresholds ξ_j, τ_j .

[0056] For each of the child nodes, the process as outlined in blocks 608-622 of FIG. 6 is recursively executed 624 for the subset of image elements directed to the respective child node. In other words, for each child node, new random test parameters are generated 608, applied 610 to the respective subset of image elements directed to the respective child node, parameters maximizing the information gain are selected 612, and the type of node (split or leaf) determined 614. If it is a leaf node then the current branch of recursion ceased. If it is a split node 620, binary tests 622 are performed to determine further subsets of image elements and another branch of recursion starts. Therefore this process recursively moves through the tree, training each node until leaf nodes are reached

at each branch. As leaf nodes are reached the process waits 626 until all the nodes in all the branches have been trained. In other examples the same functionality can be attained using alternative techniques to recursion. Once the leaf nodes have been reached for all trees in the forest the training process is complete and the process terminated 628.

5 [0057] Equation (1) is an example of a way of maximizing the confidence of the desired continuous output for all organs without intermediate voxel classification. Furthermore, the example gain formulation enables testing between different context models. For example imposing a full covariance Γ may allow correlations between all walls in all organs, with possible over-fitting consequences. In another example assuming
10 a diagonal Γ (and diagonal class covariances Λ_c) may lead to uncorrelated output predictions. In a further example Γ may be sparse but correlations between selected subgroups of classes may be enabled to capture, for example, class hierarchies or other forms of spatial context.

[0058] FIG. 7 is a flow diagram of an example method of predicting organ
15 locations in a previously unseen image. In an example a regression forest which has been trained as described in FIG. 6 can be used. An unseen image is received 700 at the organ predictor 302. An image is referred to as unseen to distinguish it from a training image which has the organ presence and/or bounding box already identified.

[0059] An image element from the unseen image is selected 702 and a trained
20 regression tree from the regression forest is also selected. The selected image is pushed 706 through the selected regression tree (in a manner similar to that described above with reference to FIG. 6) such that it is tested against the trained parameters at a node and then passed to the appropriate child in dependence on the outcome of the test. The process is repeated until the image element reaches a leaf node $l(\mathbf{v})$ with l indexing leaves across the
25 whole forest.

[0060] If it is determined 710 that there are more regression trees in the forest, then a new regression tree is selected 704, the image element is pushed through the tree until a leaf node is reached. This is repeated until it has been performed for all the regression trees 710 in the forest. In an example an image element can be pushed through the
30 plurality of trees in the regression forest in parallel. It is determined 712 whether further unanalyzed image elements are present in the unseen image, and if so another image element is selected and the process repeated. In an example the process can be repeated for every voxel in the volume $\mathbf{v} \in V$.

[0061] In an example the stored prediction at the leaf node

$p(\mathbf{d}_c | l) = N(\mathbf{d}_c; \bar{\mathbf{d}}_c, \Lambda_c)$ for class c also defines the posterior for the absolute

bounding box probability $p(\mathbf{b}_c | l) = N(\mathbf{b}_c; \bar{\mathbf{b}}_c, \Lambda_c)$ since $\bar{\mathbf{b}}_c(\mathbf{v}) = \hat{\mathbf{v}} - \bar{\mathbf{d}}_c(\mathbf{v})$. A set of

leaves $\bar{\mathcal{L}}$ can be selected. In an example $\bar{\mathcal{L}}$ may be a subset of all forest leaves. For

5 example $\bar{\mathcal{L}}$ may be the set of leaves which has the smallest uncertainty (for each class c)

and contain a specified threshold level of all test image elements. In an example the

specified threshold level is 1%. However, another level may be used. The posterior

probability for \mathbf{b}_c can also be computed. In an example the posterior probability is

$$p(\mathbf{b}_c) = \sum_{l \in \bar{\mathcal{L}}} p(\mathbf{b}_c | l) p(l).$$

10 [0062] In the above example the most confident leaves are used to predict the output. For example $p(l) = 1 / |\bar{\mathcal{L}}|$ if $l \in \bar{\mathcal{L}}$, 0 otherwise, irrespective of where in the forest they come from.

[0063] In an example, whether or not an image element is within the bounding box can be predicted 714 from the probability. In an example the organ is considered to be

15 present in the image if $p(\mathbf{b}_c = \bar{\mathbf{b}})c > \beta$. In an example $\beta = 0.5$, however any threshold value may be used. In another example the organ is considered to be present by determining the maximum probability in the overall probability distribution. The

prediction 716 for the absolute bounding box position of the c^{th} organ can given by the

expectation $\bar{\mathbf{b}}_c = \int_{\mathbf{b}_c} \mathbf{b}_c p(\mathbf{b}_c) d\mathbf{b}_c$. In an example the predictions for an image element is

20 assigned to the image element for future use.

[0064] Compared to atlas based techniques the examples methods described herein have a reduced error, and are comparably fast and more robust when computing bounding box predictions. As described above in many cases atlas based techniques become stuck in local minima which can cause inaccurate results. In some examples a local registration

25 step may be added to atlas based techniques, however this may not improve cases stuck in local minima and significantly increases the runtime of the algorithm. In addition

regression forest based techniques require significantly less memory than atlas based techniques. A regression based approach can compute the position of each wall rather than just the organ centre, thus enabling approximate extent estimation. Regression

techniques as described in the examples herein are also more accurate than classification based approaches.

[0065] FIG. 8 is a flow diagram of an example method of computing probability maps. Probability maps are an example of the type of semantic information which can be used to formulate an energy function. Estimated organ locations are received 800 from the organ predictor 302. In an example these can be the prediction for the absolute bounding box position of the c^{th} organ given by the expectation $\bar{\mathbf{b}}_c = \int_{\mathbf{b}_c} \mathbf{b}_c p(\mathbf{b}_c) d\mathbf{b}_c$ as computed using a trained regression forest algorithm as described above with reference to FIG. 7.

The outputs from the organ predictor can be transformed into a map of semantic information 802. In an example the semantic information is a probability for each image element that it lies within an organ bounding box. In an embodiment the probability map can be computed from:

$$P(B_i | x) = \frac{1}{z} \prod_{j=1}^3 \left(1 + \exp \left(\frac{x_j - u_i^j}{\sigma_i^u} \right) \right)^{-1} \left(1 + \exp \left(\frac{l_i^j - x_j}{\sigma_i^l} \right) \right)^{-1} \quad (2)$$

where B_i represents the i^{th} bounding box, $\mathbf{x} = (x_1, x_2, x_3)$ denotes image element position, u_i^j and l_i^j are the upper and lower faces of B_i in the j direction and σ_i^u and σ_i^l are the standard deviations for these coordinates as estimated by the organ predictor.

[0066] The semantic information may be output in the form of a probability map. In some embodiments a probability maps can be a volumetric probability map. In an example a plurality of probability maps can be output for each input image. For example a probability map can be output which illustrates the probability that an organ lies within one bounding box, for example the bounding box surrounding the left kidney. In an example one probability map can be output for each predicted organ bounding box. In another example an aggregate probability map can be output which illustrates the probability that an image element lies within any predicted organ bounding box.

[0067] FIG. 9 is an example of computed probability maps from an input image. In an example input image 900 may be a CT image showing the liver 902 and the right kidney 904. However, image 900 may be any appropriate image.

In an example the output of the transformer 304 is a first probability map 906 which shows the likelihood of each image element in the input image 900 being inside the liver bounding box 908 and a second probability map 910 which shows the likelihood of each

image element in the input image 900 being inside the right kidney bounding box 912. The probability maps are represented schematically using elliptical shapes to indicate regions where probability values may increase towards the center of those regions. The transformer 304 may output more than two probability maps. For example the transformer
5 may output a probability map for each bounding box prior in the image.

[0068] In other embodiments the probability maps may be aggregated probability maps which display the likelihood that an image element is inside any of the estimated bounding boxes. For example a probability map may display the probability that an image element belongs to either the liver or kidney or any other bounding box in the input image
10 900.

[0069] In an example color tables of the probability maps are scaled according to the likelihood that an image element is within a bounding box. For example in an example where there are 256 colors in a color table, an image element with a probability of being inside the bounding box of 1.0 is set as color 256 and an image element with a probability
15 of being inside the bounding box of 0 is set as color 1. In other examples the colors may be scaled differently, for example logarithmically or according to a gamma scaling.

[0070] Although the probability maps as described in FIG. 9 are displayed to the user, this is not essential. In other embodiments the probability maps may be computed as described in FIG. 8 without displaying the output of the transformer to the user. For
20 example the computed probability maps or other semantic data may be held temporarily in cache memory, stored at a data store or otherwise retained without displaying the information to the user.

[0071] FIG. 10 is a flow diagram of an example method of image registration. Semantic data is received 1000 from the transformer 304 at the optimizer 308. In an
25 embodiment the problem of how to register the reference image and the target image is an energy optimization problem. In an embodiment the energy optimization problem comprises finding a global point of inflection, for example the minimum or maximum of an energy function. For example the energy function 206 described above with reference to FIG. 2. An example method of computing the energy optimization 1002 is to formulate
30 the energy optimization problem as an energy minimization task.

[0072] In order to compute the energy optimization function the mutual information may be computed 1004 between the reference image and the target image. The mutual information is a quantity that measures the mutual dependence of two

variables, for example an image element in reference image I and an image element in target image J . In an example high mutual information indicates a large reduction in uncertainty; low mutual information indicates a small reduction; and zero mutual information between two random variables means the variables are independent.

5 [0073] In an example x is a random variable over coordinate locations in the reference image. In some embodiments samples are drawn from x in order to approximate the mutual information. In an example the mutual information may be defined as

$$MI(I(x), J(T(x))) \equiv h(I(x)) + h(J(T(x))) - h(I(x), J(T(x)))$$

where $h(I(x))$ is the marginal entropy of the reference image and $h(J(T(x)))$ is the marginal entropy of the part of the test image into which the reference image projects. $h(I(x), J(T(x)))$ is the joint entropy of the reference image and the target image. In an embodiment $h(\square)$ is the entropy of a random variable and is defined as

$$h(x) = -\int p(x, y) \ln p(x, y) dx dy.$$

[0074] In some embodiments a Kullback-Leibler (KL) divergence may be computed 1006. The KL divergence is a non symmetric measure of the distance between two probability distributions. An example of computing the probability distributions is described above with reference to box 802 of FIG. 8 The KL divergence can be given by

$$KL(I(x), J(T(x))) = \int_{\Omega} p(x) \sum_l P(p_i^l | x) \ln \left(\frac{P(B_i^I | x)}{P(B_i^J | T(x))} \right) dx$$

[0075] In an embodiment the energy optimization problem is an energy minimization task which may be computed from the mutual information and the KL divergence. For example the energy minimization problem may be:

$$T^* = \arg \min_T E(T), E(T) = -MI(I(x), J(T(x))) + KL(I(x), J(T(x)))$$

where I is the target image, J is the moving image, T is the spatial transformation, MI is the mutual information, and KL is the Kullback-Leibler divergence and the point probability distribution $p(x)$ is assumed uniform over the image domain Ω . In other examples the energy optimization problem may be an energy maximization task.

[0076] In other embodiments other methods may be used to compute an energy optimization. For example other entropy measures may be used. An example of an alternative entropy measure is the Shannon entropy.

[0077] Global optimization of the energy function $E(T)$ enables a linear spatial transformation 1008 used to register the images to be computed and the target image to be registered 1010 to the reference image. In other embodiments a non-linear spatial transformation may be computed.

5 [0078] Fig. 11 is an example of a registered image. The image 1100 may be a CT scan or other appropriate image. In this example the image shows the reference image of the right kidney 1102 with the registered target image of the right kidney overlaid as a contour map 1104. In an example the contours may be intensity contours, for example contours at 90, 70 and 60 percent of the intensity of the reference image. In other
10 examples the registered target image may be overlaid as a partially transparent image or displayed sequentially.

[0079] FIG. 12 illustrates various components of an exemplary computing-based device 1200 which may be implemented as any form of a computing and/or electronic device, and in which embodiments of an image registration system may be implemented.

15 [0080] The computing-based device 1200 comprises one or more inputs 1202 which are of any suitable type for receiving media content, Internet Protocol (IP) input, volumetric image data, input from medical devices, input from image capture devices, input from users, input from other computing-based devices. The device also comprises communication interface 1204 which may be used to enable the device to communicate
20 with other devices over a communications network.

[0081] Computing-based device 1200 also comprises one or more processors 1206 which may be microprocessors, controllers or any other suitable type of processors for processing computing executable instructions to control the operation of the device in order to perform image registration. The processors 1206 may comprise one or more
25 specialist image processing processors such as graphics processing units. In some examples, for example where a system on a chip architecture is used, the processors 1206 may include one or more fixed function blocks (also referred to as accelerators) which implement a part of a method of image registration in hardware (rather than software or firmware). Platform software comprising an operating system 1208 or any other suitable
30 platform software may be provided at the computing-based device to enable application software 1210 to be executed on the device.

[0082] Further software that can be provided at the computing-based device 1200 includes tree training logic 1212 (which implements the techniques described for training

a regression tree, or other appropriate tree training techniques), image analysis logic 1214 (which implements techniques of passing an unseen image through a regression tree as described herein), image transform logic 1216 (which computes semantic information as described herein) and image optimization logic 1218 (which solves the registration
5 problem). A data store 1220 is provided to store data such as the training parameters, probability distributions, and analysis results.

[0083] The computer executable instructions may be provided using any computer-readable media that is accessible by computing based device 1200. Computer-readable media may include, for example, computer storage media such as memory 1222
10 and communications media. Computer storage media, such as memory 1222, includes volatile and non-volatile, removable and non-removable media implemented in any method or technology for storage of information such as computer readable instructions, data structures, program modules or other data. Computer storage media includes, but is not limited to, RAM, ROM, EPROM, EEPROM, flash memory or other memory
15 technology, CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other non-transmission medium that can be used to store information for access by a computing device. In contrast, communication media may embody computer readable instructions, data structures, program modules, or other data in a modulated data signal,
20 such as a carrier wave, or other transport mechanism. As defined herein, computer storage media does not include communication media. Although the computer storage media (memory 1222) is shown within the computing-based device 1200 it will be appreciated that the storage may be distributed or located remotely and accessed via a network or other communication link (e.g. using communication interface 1204).

[0084] An output 1224 is also provided such as an audio and/or video output to a display system integral with or in communication with the computing-based device. The display system may provide a graphical user interface, or other user interface of any
25 suitable type although this is not essential. In some examples the display system comprises a touch sensitive display screen.

[0085] The term 'computer' is used herein to refer to any device with processing capability such that it can execute instructions. Those skilled in the art will realize that such processing capabilities are incorporated into many different devices and therefore the term 'computer' includes PCs, servers, mobile telephones, personal digital assistants and
30 many other devices.

[0086] The methods described herein may be performed by software in machine readable form on a tangible storage medium e.g. in the form of a computer program comprising computer program code means adapted to perform all the steps of any of the methods described herein when the program is run on a computer and where the computer
5 program may be embodied on a computer readable medium. Examples of tangible (or non-transitory) storage media include disks, thumb drives, memory etc and do not include propagated signals. The software can be suitable for execution on a parallel processor or a serial processor such that the method steps may be carried out in any suitable order, or simultaneously.

10 [0087] This acknowledges that software can be a valuable, separately tradable commodity. It is intended to encompass software, which runs on or controls “dumb” or standard hardware, to carry out the desired functions. It is also intended to encompass software which “describes” or defines the configuration of hardware, such as HDL (hardware description language) software, as is used for designing silicon chips, or for
15 configuring universal programmable chips, to carry out desired functions.

[0088] Those skilled in the art will realize that storage devices utilized to store program instructions can be distributed across a network. For example, a remote computer may store an example of the process described as software. A local or terminal computer may access the remote computer and download a part or all of the software to run the
20 program. Alternatively, the local computer may download pieces of the software as needed, or execute some software instructions at the local terminal and some at the remote computer (or computer network). Those skilled in the art will also realize that by utilizing conventional techniques known to those skilled in the art that all, or a portion of the software instructions may be carried out by a dedicated circuit, such as a DSP,
25 programmable logic array, or the like.

[0089] Any range or device value given herein may be extended or altered without losing the effect sought, as will be apparent to the skilled person.

[0090] It will be understood that the benefits and advantages described above may relate to one embodiment or may relate to several embodiments. The embodiments are not
30 limited to those that solve any or all of the stated problems or those that have any or all of the stated benefits and advantages. It will further be understood that reference to 'an' item refers to one or more of those items.

[0091] The steps of the methods described herein may be carried out in any suitable order, or simultaneously where appropriate. Additionally, individual blocks may

be deleted from any of the methods without departing from the spirit and scope of the subject matter described herein. Aspects of any of the examples described above may be combined with aspects of any of the other examples described to form further examples without losing the effect sought.

5 [0092] The term 'comprising' is used herein to mean including the method blocks or elements identified, but that such blocks or elements do not comprise an exclusive list and a method or apparatus may contain additional blocks or elements.

[0093] It will be understood that the above description of a preferred embodiment is given by way of example only and that various modifications may be made by those skilled in the art. The above specification, examples and data provide a complete
10 description of the structure and use of exemplary embodiments of the invention. Although various embodiments of the invention have been described above with a certain degree of particularity, or with reference to one or more individual embodiments, those skilled in the art could make numerous alterations to the disclosed embodiments without departing from
15 the spirit or scope of this invention.

CLAIMS

1. A computer-implemented method of automatic image registration comprising:
 - receiving a first image and a second image both images being of objects and with
5 at least part of one object being common to both images;
 - for each of the first and second images computing a probability map comprising, for each image element, a probability that the image element is of a specified object;
 - finding a mapping to register the first and second images by optimizing an energy
10 function which is a function of the intensities of the first and second images and also of the probability maps.
2. A method as claimed in claim 1 wherein the first and second images are medical images.
3. A method as claimed in claim 1 or claim 2 wherein the first and second images are of different modalities.
- 15 4. A method as claimed in any preceding claim wherein the first and second images are of three or higher dimensions.
5. A method as claimed in claim 1 wherein computing the probability map comprises, for each image element, computing a plurality of probabilities such that there is one probability for each of a plurality of specified objects associated with each image
20 element.
6. A method as claimed in any preceding claim wherein computing the probability map comprises clustering image elements together based on their appearance, spatial context and their confidence in predictions of bounding boxes of at least one specified object.
- 25 7. A method as claimed in any preceding claim wherein computing the probability map comprises using a regression forest comprising a plurality of regression trees each having been trained to predict locations of the specified object.
8. A method as claimed in any preceding claim wherein computing the probability map comprises using a classification forest comprising a plurality of
30 classification trees having been trained to classify image elements into specified categories.
9. A method as claimed in any preceding claim wherein finding the mapping comprises globally optimizing the energy function.

10. An automatic image registration system comprising:
- an input arranged to receive a first image and a second image both images being of objects and with at least part of one object being common to both images;
 - a processor arranged to compute semantic information for each of the first and
5 second images, the semantic information comprising information about the types of objects in the images and the certainty of that information;
 - the processor further arranged to find a mapping to register the first and second images which takes into account the intensities of the image elements of the images and also takes into account the semantic information in a manner which is weighted by the
10 certainty of that semantic information.

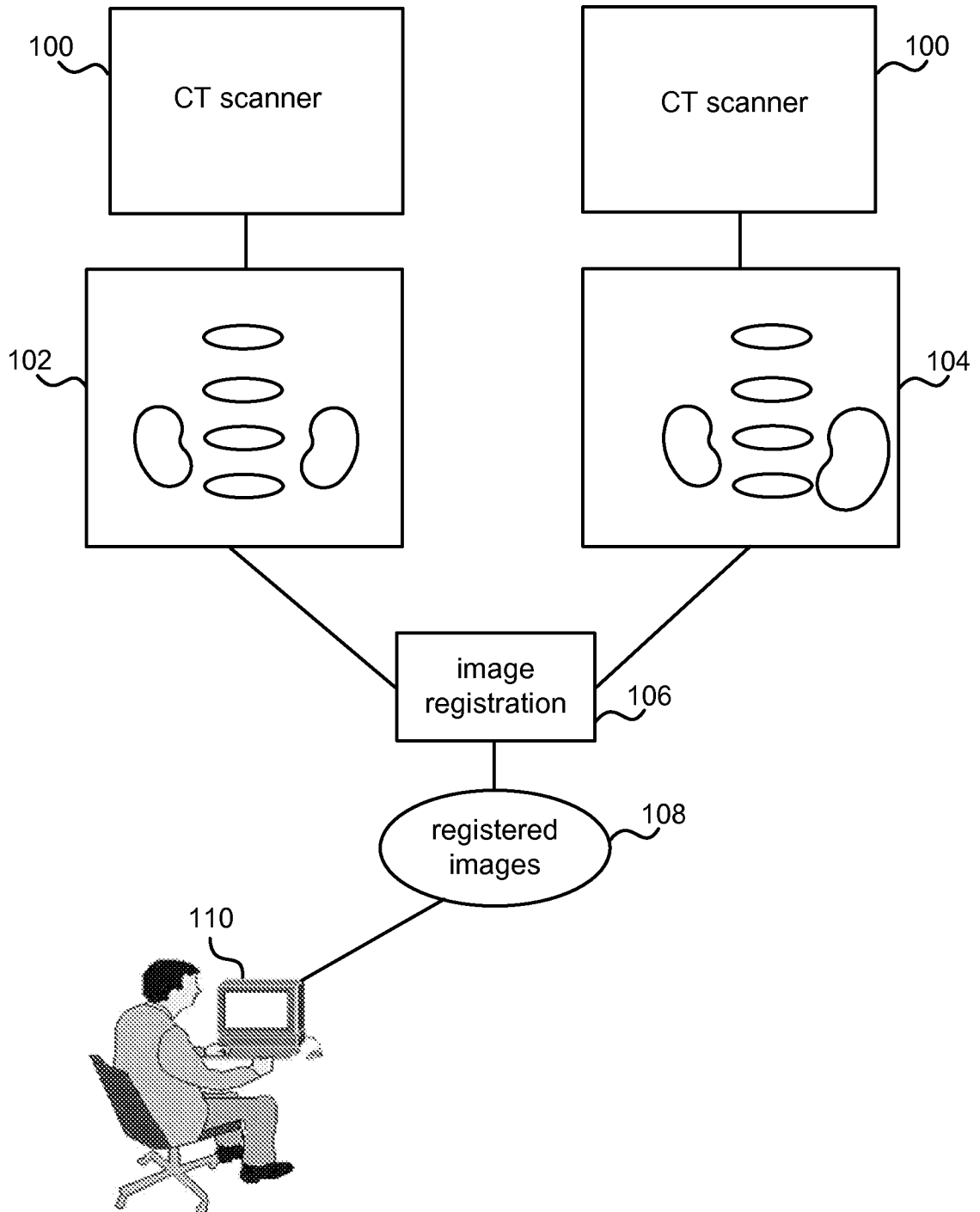


FIG. 1

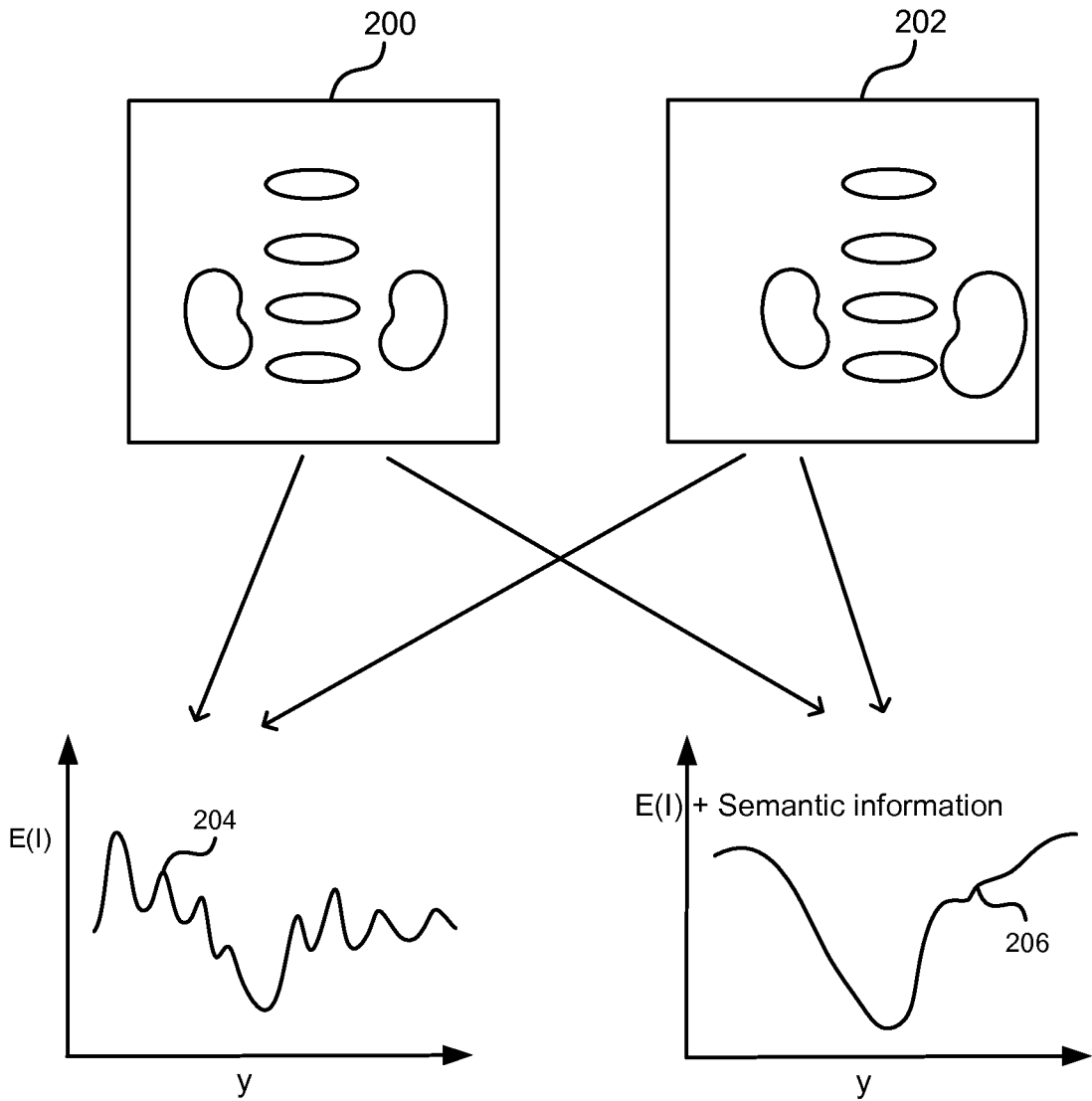


FIG. 2

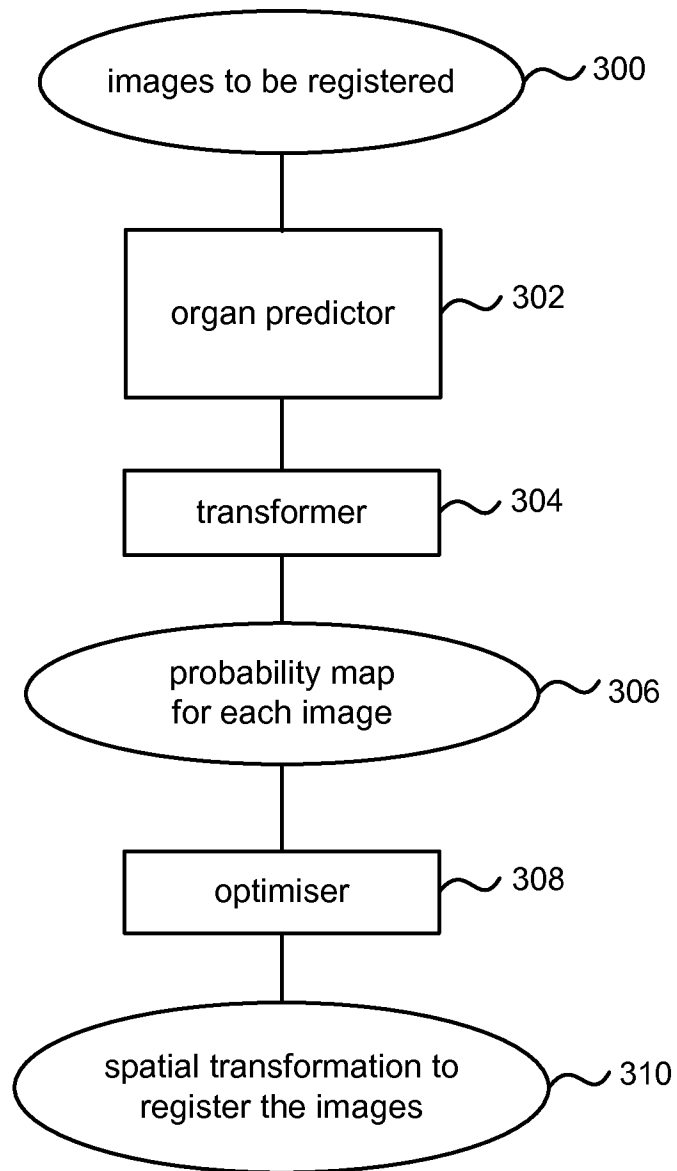


FIG. 3

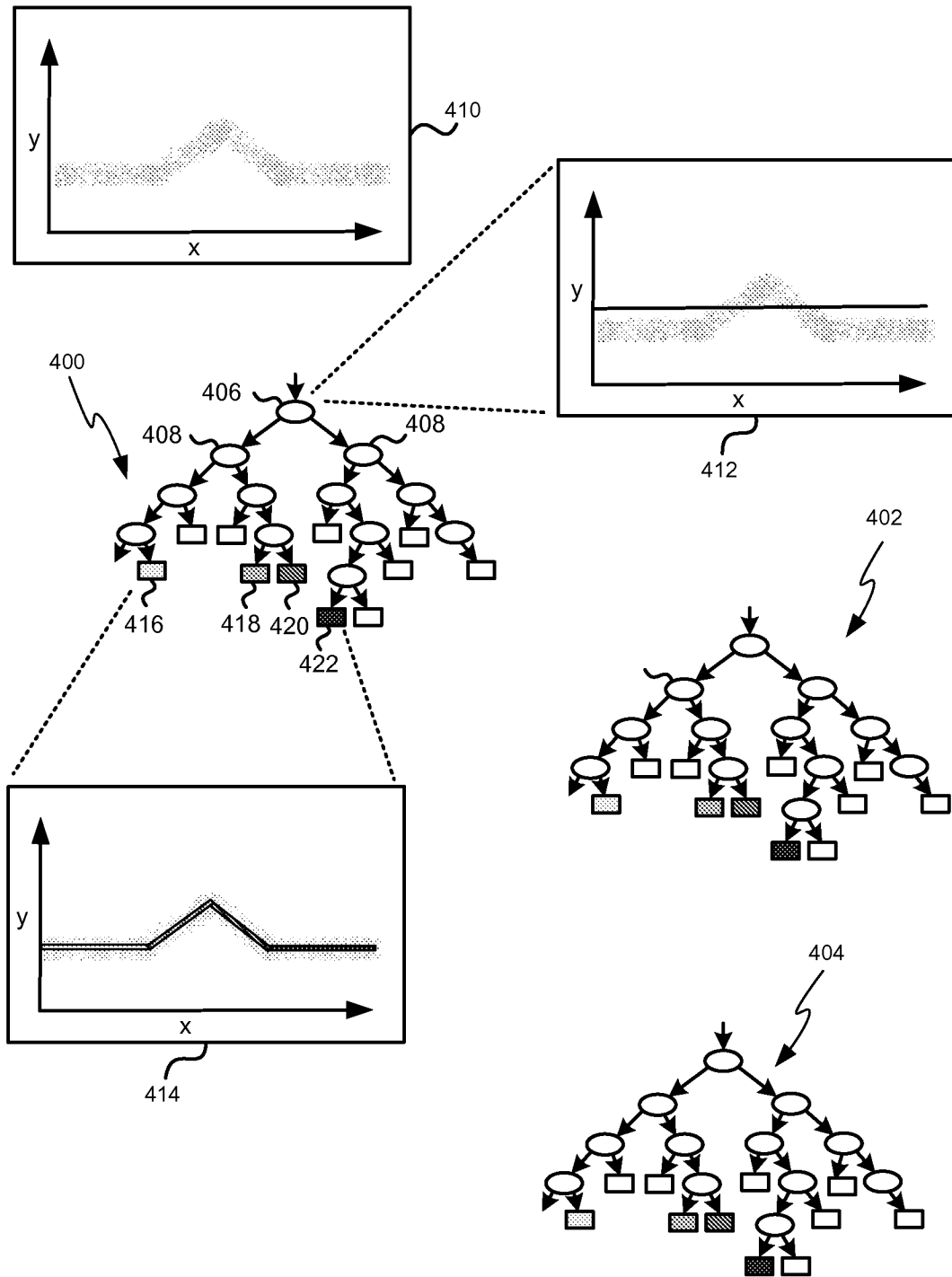


FIG. 4

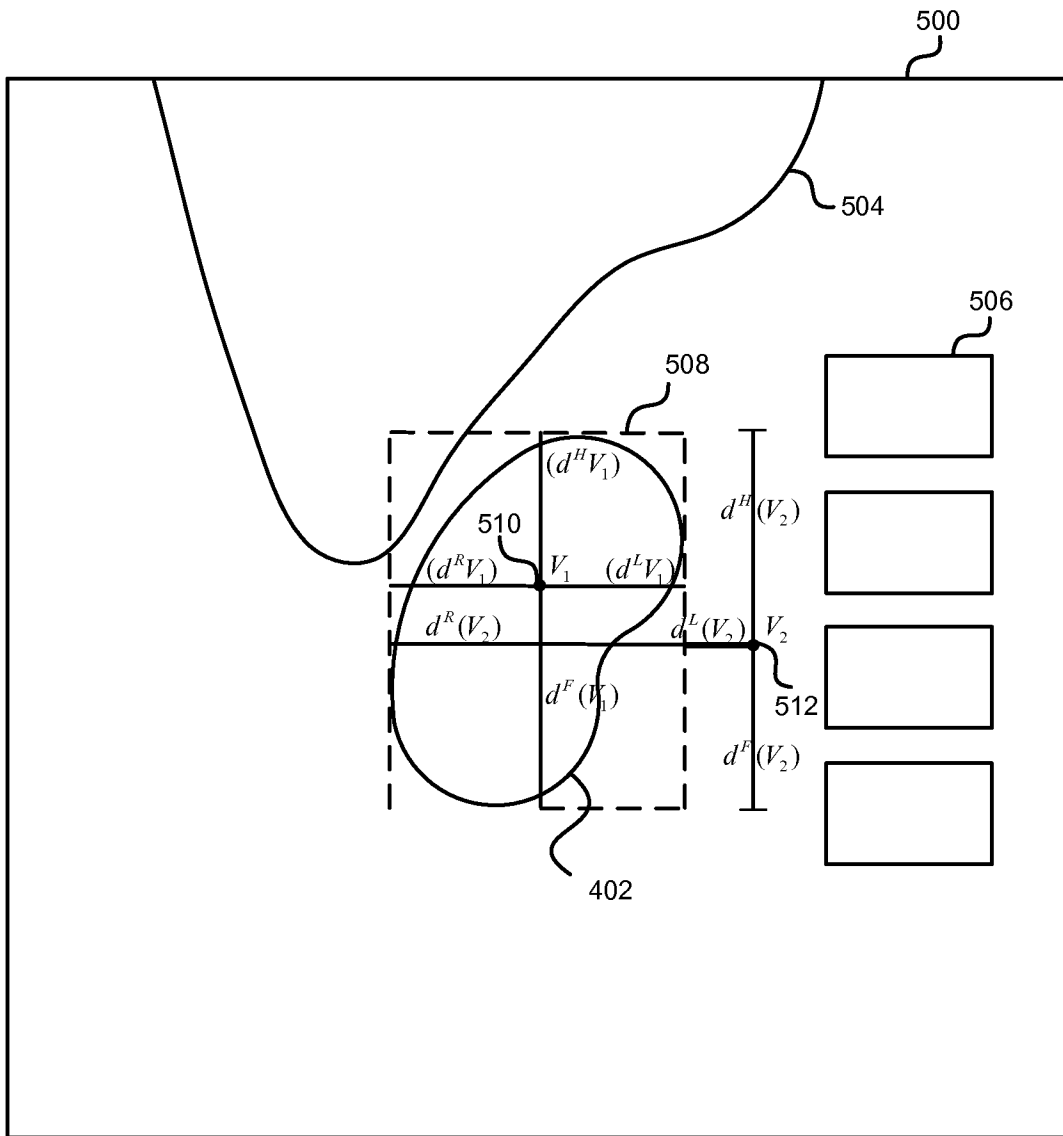


FIG. 5

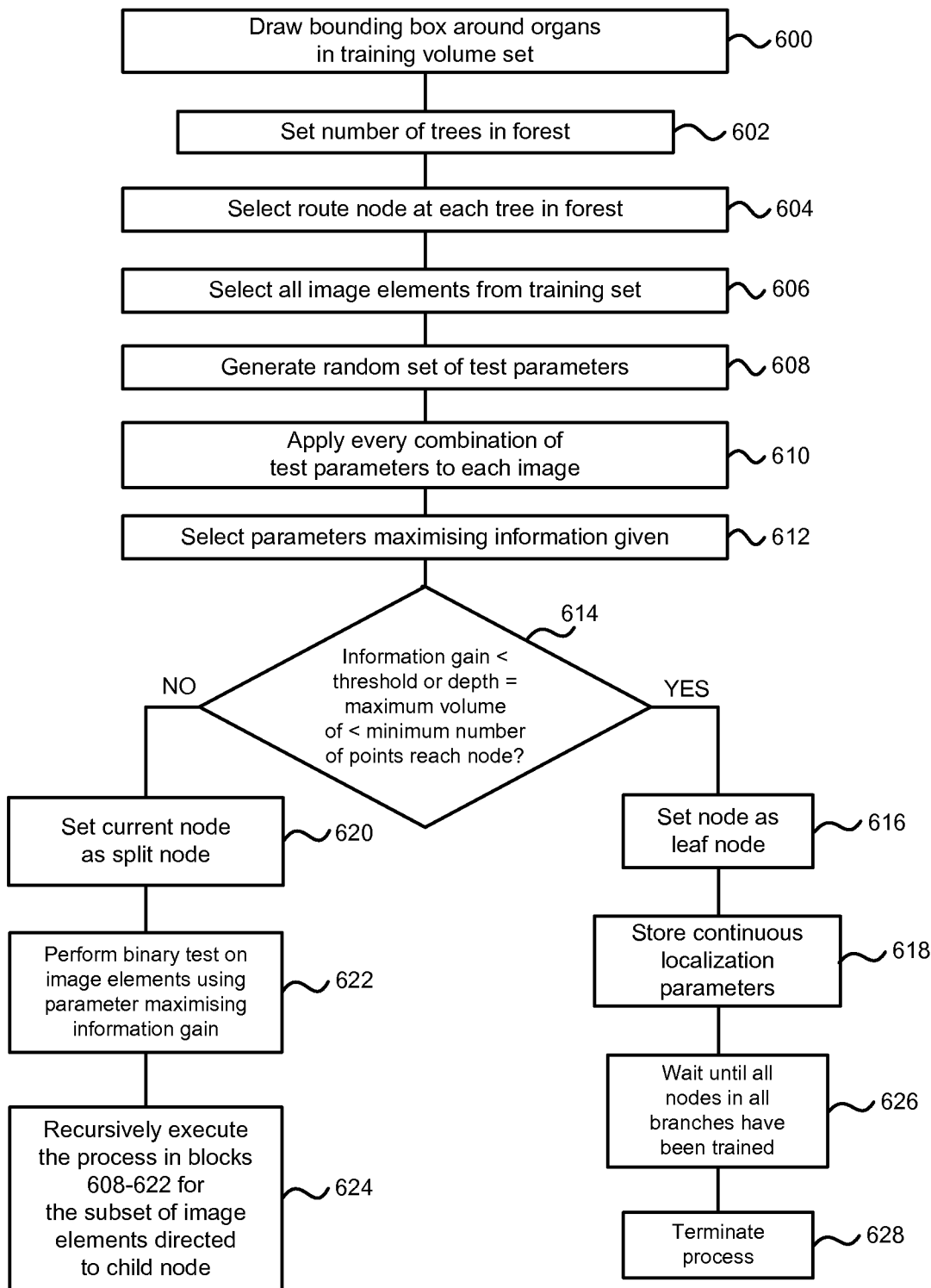


FIG. 6

7/12

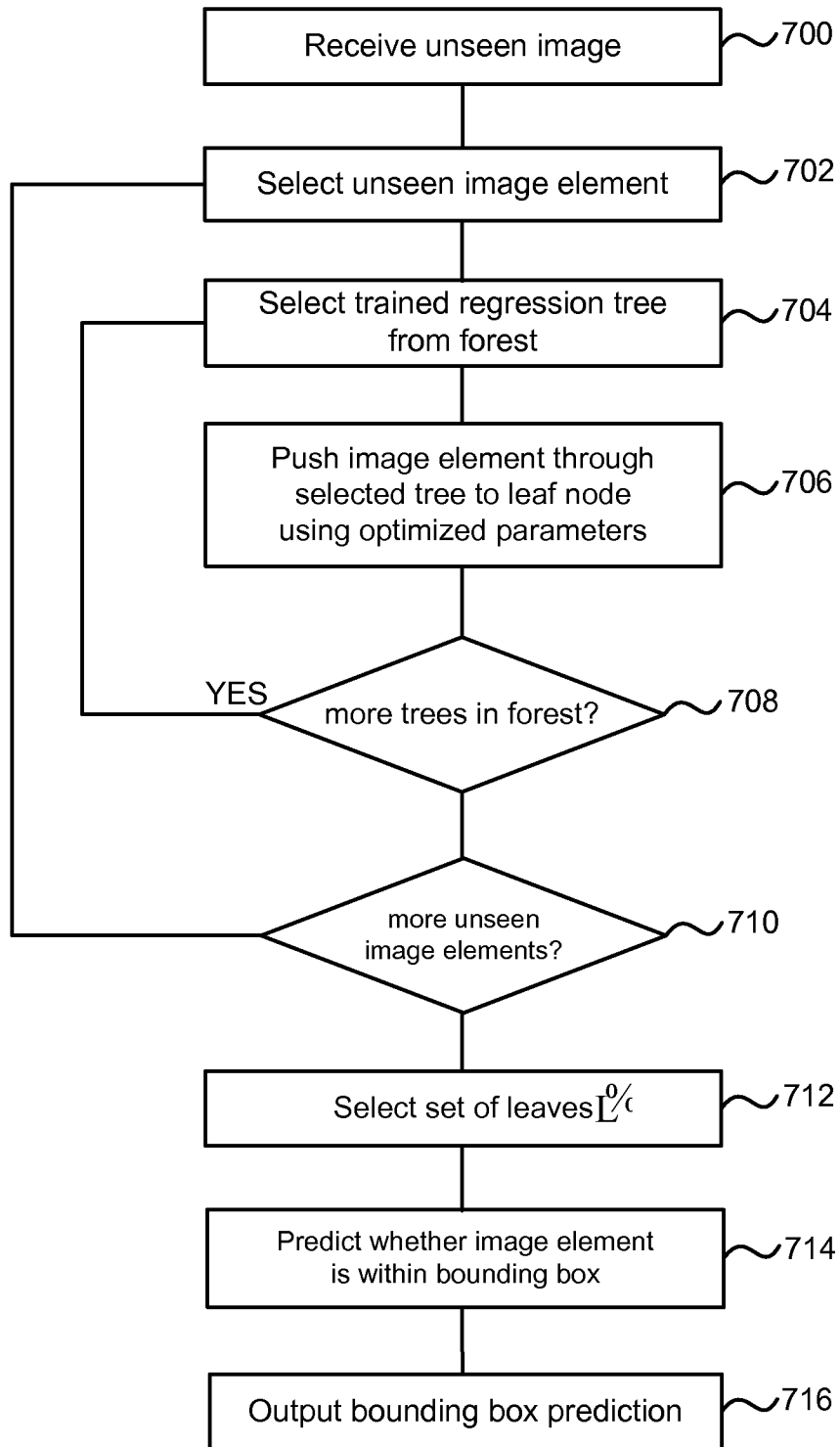


FIG. 7

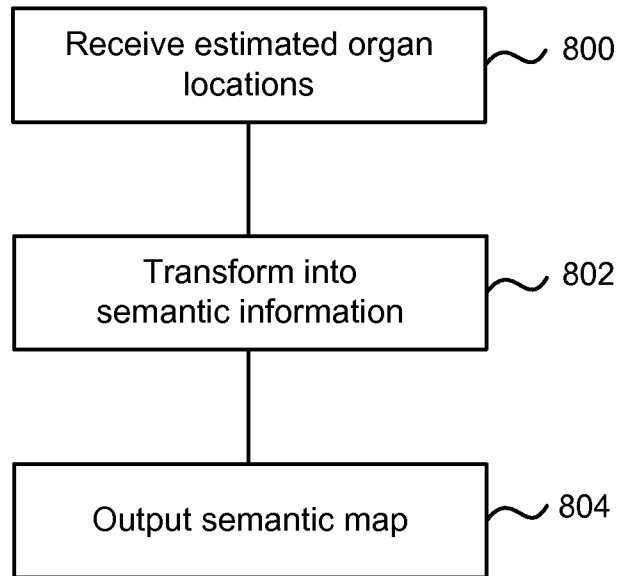


FIG. 8

9/12

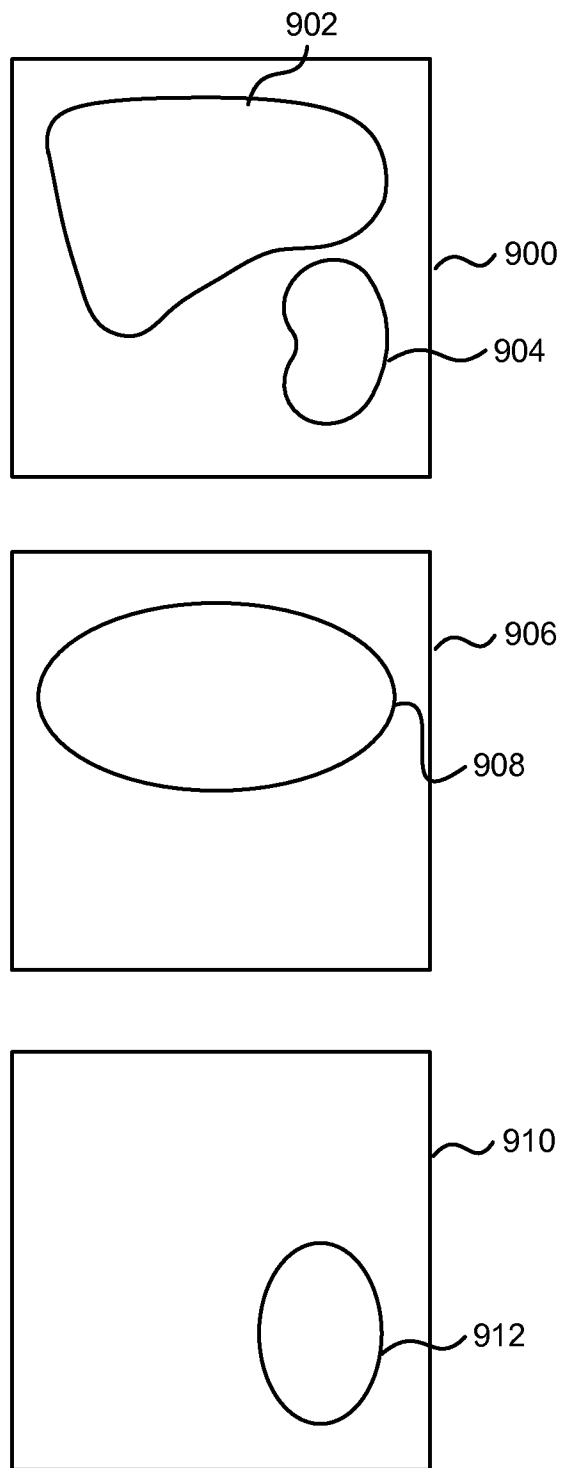


FIG. 9

10/12

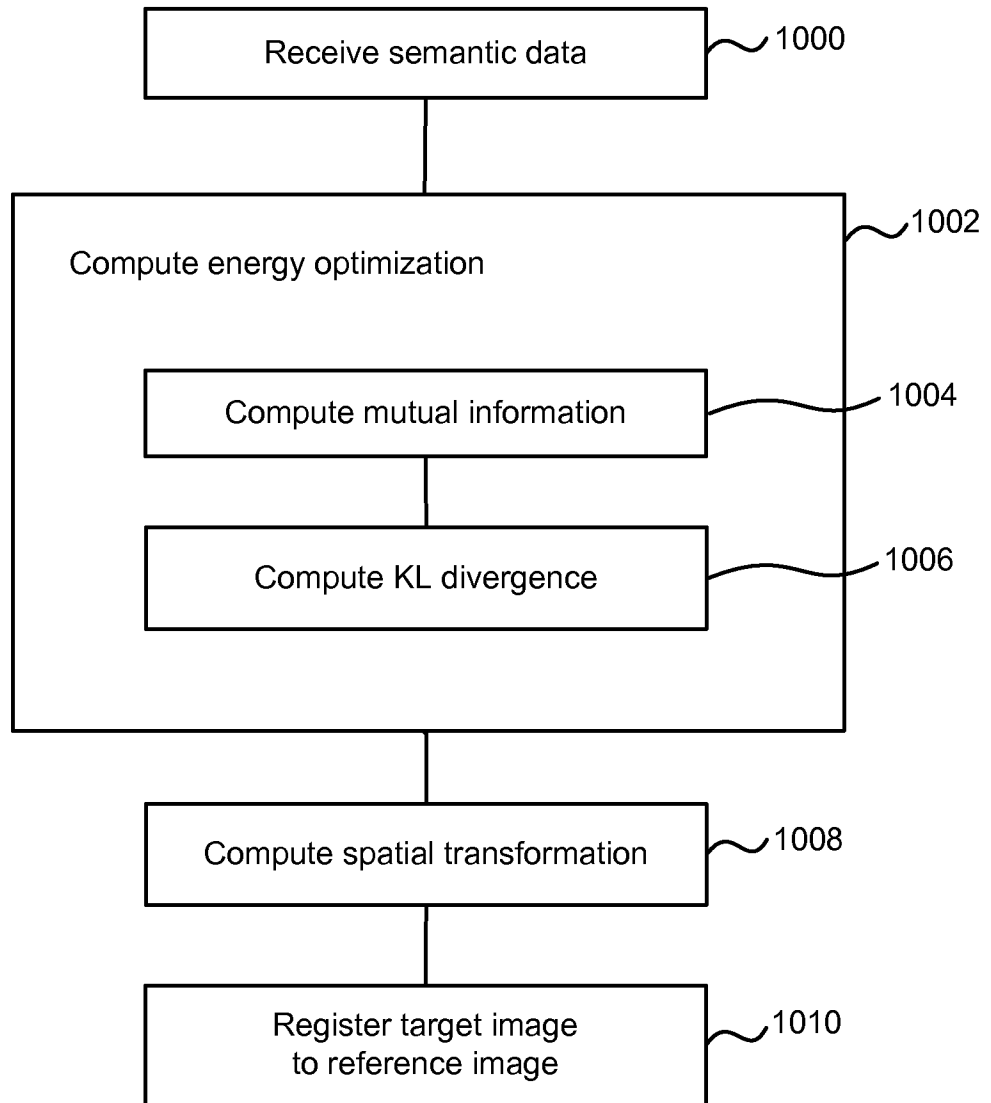


FIG. 10

11/12

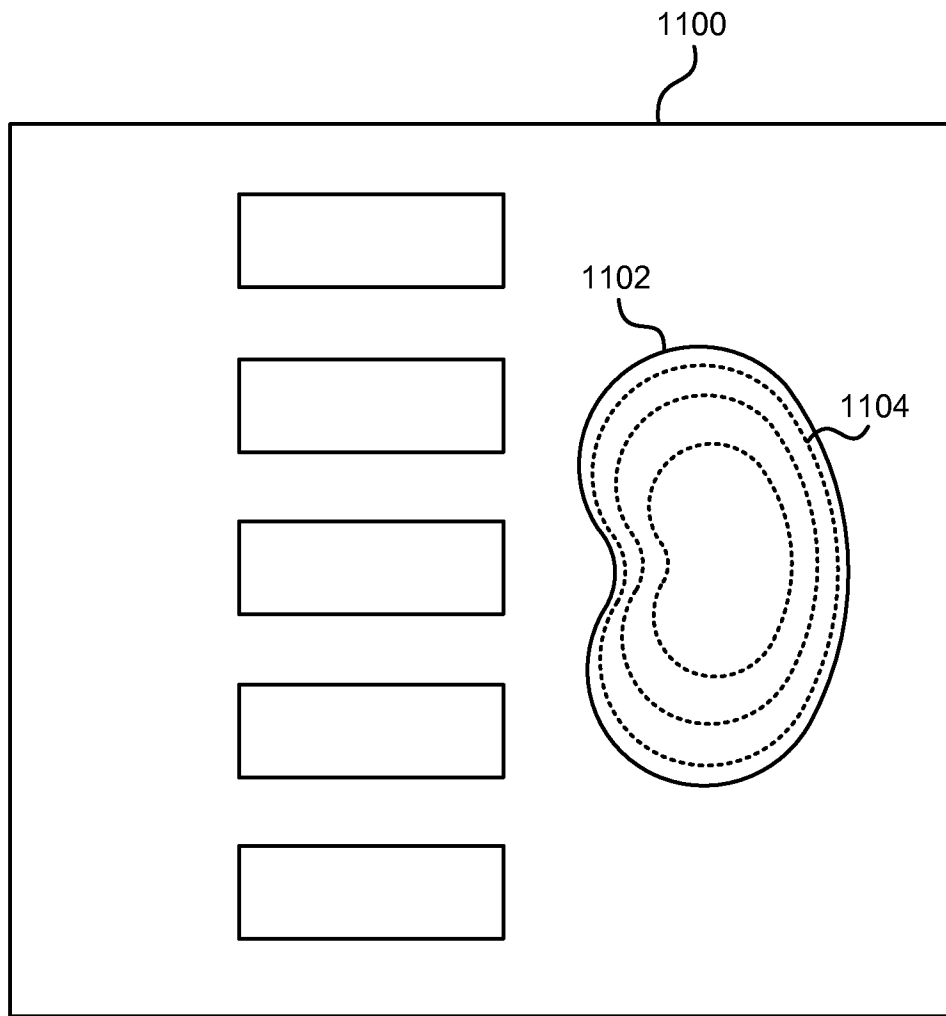


FIG. 11

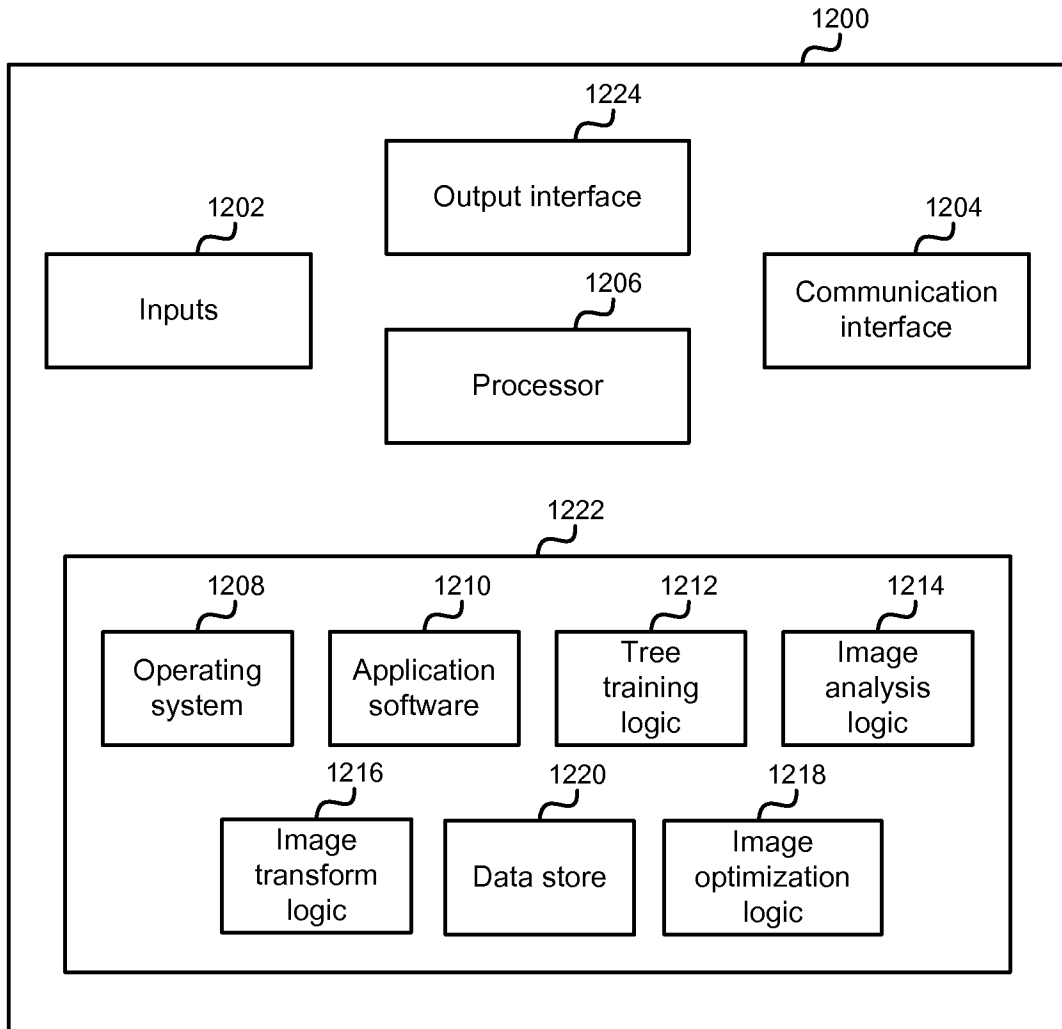


FIG. 12