

(12) Indian Patent Application

(21) Application Number: 202241013469

(22) Filing Date: 11/03/2022 (43) Publication Date: 15/09/2023

(71) Applicant(s): L&T TECHNOLOGY SERVICES LIMITED

(72) Inventor(s): Vasudevan, Shriram Kris

(51) International Classifications: G06K 9/62 G06K 9/46 G01B 11/02 G06T 7/12 G06T 7/62

(54) Title: METHOD AND SYSTEM FOR DIMENSIONING AN OBJECT

(57) Abstract: The present disclosure describes technique for dimensioning an object. The technique includes implementing region identification model on image of the object to identify ROI in the image and a bounding box for the ROI. The bounding box comprises first set of pixel coordinates corresponding to vertices of the bounding box and first overlap value indicating degree of overlap of the bounding box upon the object. The technique includes applying auxiliary classifiers upon the region identification model to adjust the bounding box to new bounding box comprising second set of pixel coordinates and second overlap value greater than the first overlap value. The technique includes determining image dimensions of the object based on the second set of pixel coordinates of the second bounding box. The real dimensions of the object is determined based on the image dimensions of the object, and ratio obtained between real and image dimensions of reference object.

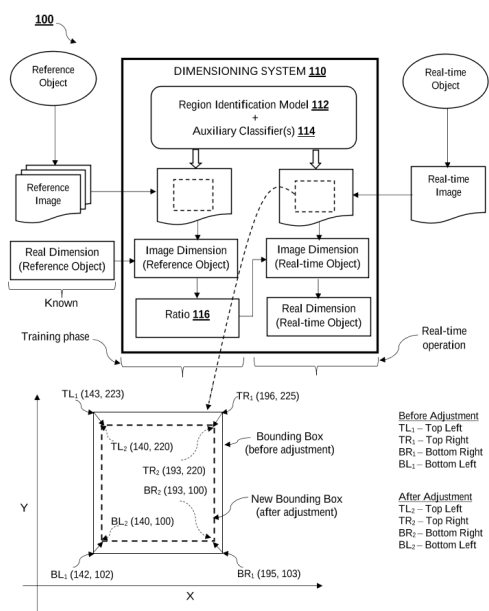


Figure 1 Robin Koshy Varghese (INPA No: 3705) Head, IPR Dept.,

FORM 2

THE PATENTS ACT 1970
(39 OF 1970)
&
The Patent Rules, 2003
Complete Specification
(See Section 10 and Rule 13)

1. TITLE OF THE INVENTION

METHOD AND SYSTEM FOR DIMENSIONING AN OBJECT

2. APPLICANT(S)

(a) NAME : **L&T TECHNOLOGY SERVICES LIMITED**
(b) NATIONALITY : **INDIAN**
(c) ADDRESS : **DLF IT SEZ Park, 2nd Floor – Block 3,
1/124, Mount Poonamallee Road, Ramapuram,
Chennai – 600 089, INDIA.**

3. PREAMBLE TO THE DESCRIPTION

COMPLETE

The following specification particularly describes the invention and the manner in which it is to be performed.

DESCRIPTION

TECHNICAL FIELD

[0001] The present disclosure generally relates to the field of object dimensioning technique, and more particularly to an implementation of deep learning techniques for object dimensioning.

BACKGROUND OF THE INVENTION

[0002] Dimensioning an object is about measuring length, width and height of the objects. Though it sounds simple, however it becomes complex when the object to be dimensioned comes in different shapes and sizes. Dimensioning is mainly implemented in industries involved in manufacturing products/components of different shapes and sizes. In such industries, accuracy is an important factor while dimensioning the objects.

[0003] Various techniques are used for improving the accuracy while dimensioning the objects. One of the widely used technique is image processing, in which, images of object are captured and then analyzed to determine the dimension. Though such imaging techniques provide a certain level of accuracy, however they lacks precision when the shape and size of the objects becomes complex.

[0004] Nowadays, various Machine Learning (ML) and deep learning techniques have been developed for dimensioning the objects. These techniques use deep learning models for dimensioning the objects by processing the images. Initially the deep learning models are trained over a plurality of reference images of reference objects. The accuracy of the deep learning model for dimensioning the objects in real-time depends on the type of reference images with which they are trained. In many instances when the object to be dimension in the real-time significantly varies from the reference objects in terms of their shape and size, such deep learning models fail to precisely dimension the object.

[0005] Thus, there exists a need for efficient techniques that can precisely dimension the objects irrespective of their shapes and sizes.

SUMMARY OF THE INVENTION

[0006] One or more shortcomings discussed above are overcome, and additional advantages are provided by the present disclosure. Additional features and advantages are realized through the techniques of the present disclosure. Other embodiments and aspects of the disclosure are described in detail herein and are considered a part of the disclosure.

[0007] An object of the present disclosure is to provide efficient techniques for improving existing neural network based image dimensioning technique.

[0008] Another objective of the present disclosure is to provide efficient techniques for precisely dimensioning an object using the improved image dimensioning technique.

[0009] The above stated objects as well as other objects, features, and advantages of the present disclosure will become clear to those skilled in the art upon review of the following description, the attached drawings, and the appended claims.

[0010] In a non-limiting embodiment, a method for dimensioning an object is disclosed. The method may comprise implementing a region identification model on an image of the object to be dimensioned. The implementing of the region identification model may further comprise identifying a region of interest (ROI) in the image of the object, and determining a bounding box for the ROI in the image. The bounding box comprises a first set of pixel coordinates corresponding to vertices of the bounding box and a first overlap value indicating a degree of overlap of the bounding box upon the object. The method may further comprise applying one or more auxiliary classifiers upon the region identification model in order to enable the region identification model to adjust the bounding box to a new bounding box such that new bounding box comprises a second set of pixel coordinates and a second overlap value. The second overlap value is greater than the first overlap value, thereby indicating that the new bounding box having higher degree of overlap compared to the bounding box before adjustment. The method may further comprise determining an image dimensions of the object based on the second set of pixel coordinates of the second bounding box. Further, the method may comprise determining a real dimensions of the object based on the image dimensions of the object, and a ratio obtained between a real dimensions and an image dimensions of a reference object.

[0011] In a non-limiting embodiment, a dimensioning system for dimensioning an object is disclosed. The dimensioning system comprises a memory and at least one processor operatively coupled to the memory. The at least one processor is configured to implement a region identification model on an image of the object to be dimensioned. The implementing of the region identification model comprises identifying a region of interest (ROI) in the image of the object, and determining a bounding box for the ROI in the image. The bounding box comprises a first set of pixel coordinates corresponding to vertices of the bounding box and a first overlap value indicating a degree of overlap of the bounding box upon the object. The at least one processor is further configured to apply one or more auxiliary classifiers upon the region identification model in order to enable the region identification model to adjust the bounding box to a new bounding box such that new bounding box comprises a second set of pixel coordinates and a second overlap value. The second overlap value is greater than the first overlap value, thereby indicating that the new bounding box having higher degree of overlap compared to the bounding box before adjustment. The at least one processor is further configured to determine an image dimensions of the object based on the second set of pixel coordinates of the second bounding box. Further, at least one processor is configured to determine a real dimensions of the object based on the image dimensions of the object, and a ratio obtained between a real dimensions and an image dimensions of a reference object.

[0012] The foregoing summary is illustrative only and is not intended to be in any way limiting. In addition to the illustrative aspects, embodiments, and features described above, further aspects,

embodiments, and features will become apparent by reference to the drawings and the following detailed description.

BRIEF DESCRIPTION OF DRAWINGS

[0013] Further aspects and advantages of the present disclosure will be readily understood from the following detailed description with reference to the accompanying drawings. Reference numerals have been used to refer to identical or functionally similar elements. The figures together with a detailed description below, are incorporated in and form part of the specification, and serve to further illustrate the embodiments and explain various principles and advantages, in accordance with the present disclosure wherein:

[0014] **Figure 1** shows an exemplary environment of a dimensioning system **100** for dimensioning an object, in accordance with some embodiments of the present disclosure;

[0015] **Figure 2** shows a detailed block diagram **200** of the dimensioning system **100** illustrated in **Figure 1**, in accordance with some embodiments of the present disclosure;

[0016] **Figure 3** shows a process flow diagram for dimensioning an object, in accordance with some embodiments of the present disclosure: and

[0017] **Figure 4** shows a process flow diagram for determining a ratio, in accordance with some embodiments of the present disclosure.

[0018] It should be appreciated by those skilled in the art that any block diagrams herein represent conceptual views of the illustrative systems embodying the principles of the present disclosure. Similarly, it will be appreciated that any flowcharts, flow diagrams, state transition diagrams, pseudo code, and the like represent various processes which may be substantially represented in computer readable medium and executed by a computer or processor, whether or not such computer or processor is explicitly shown.

DETAILED DESCRIPTION

[0019] In the present document, the word “exemplary” is used herein to mean “serving as an example, instance, or illustration.” Any embodiment or implementation of the present disclosure described herein as “exemplary” is not necessarily to be construed as preferred or advantageous over other embodiments.

[0020] While the disclosure is susceptible to various modifications and alternative forms, specific embodiments thereof have been shown by way of example in the drawings and will be described in detail below. It should be understood, however, that it is not intended to limit the disclosure to the particular form disclosed, but on the contrary, the disclosure is to cover all modifications, equivalents, and alternatives falling within the spirit and the scope of the disclosure.

[0021] The terms “comprise(s)”, “comprising”, “include(s)”, or any other variations thereof, are intended to cover a non-exclusive inclusion, such that a setup, device, apparatus, system, or method that comprises a list of components or steps does not include only those components or steps but may include other components or steps not expressly listed or inherent to such setup or device or apparatus or system or method. In other words, one or more elements in a device or system or apparatus preceded by “comprises... a” does not, without more constraints, preclude the existence of other elements or additional elements in the system.

[0022] The terms like “at least one” and “one or more” may be used interchangeably throughout the description. The terms like “a plurality of” and “multiple” may be used interchangeably throughout the description.

[0023] In the following detailed description of the embodiments of the disclosure, reference is made to the accompanying drawings that form a part hereof, and in which are shown by way of illustration of specific embodiments in which the disclosure may be practiced. These embodiments are described in sufficient detail to enable those skilled in the art to practice the disclosure, and it is to be understood that other embodiments may be utilized and that changes may be made without departing from the scope of the present disclosure. The following description is, therefore, not to be taken in a limiting sense. In the following description, well known functions or constructions are not described in detail since they would obscure the description with unnecessary detail.

[0024] Disclosed herein is a dimensioning technique for dimensioning objects of different shape and size. Dimensioning of the objects are done in various industries mainly involved in manufacturing of products, for example automobile industry, original equipment manufacturer (OEM), digital manufacturing industry, medical component industry and the like. The products or components are manufactured in different shapes and sizes, and they may include for example, but not limited to, nuts and bolts, screws, keys, logos, electronic components etc. Precision or an accuracy plays a vital role in determining quality of these products. Even a minor dimension error in millimeters (mm) in these components can impact entire system in which they are intended to be implemented. This further leads to rejection of these products during quality check process and may therefore impact entire business operation.

[0025] Conventional dimensioning techniques implements image processing along machine learning model like neural network. However, the implementation of known techniques are dependent upon reference data i.e., prestored dimensions of reference objects. The neural network are made up of multiple neurons which acts as processing/computing units and are arranged in different layers like input layer, hidden layer and output layer. These layers are interconnected layers and perform data processing in such a manner that one layer takes an input from one or more neurons, process the input, and passes an output to one or more neurons of next layer. Based on the correctness of the output, the neurons pass the feedback to the one or more neurons in backward direction. This entire process helps the neural network to learn about any task/operation over the time. The dimensioning technique of the

present disclosure takes this opportunity and leverages the functioning of the neural network. That is, the disclosed dimensioning technique implements a region identification model such as Region Proposal Network (RPN) and further leverages its functioning by applying auxiliary classifiers which are convolution neural network classifiers (CNN classifiers). The combination of RPN and auxiliary classifiers provides better precision in dimensioning the objects, which will be now explained in detail in the specification while referring to drawings.

[0026] Referring now to **Figure 1**, which illustrates an exemplary environment **100** for dimensioning an object, in accordance with some embodiments of the present disclosure. The environment **100** comprises a dimensioning system **110** which is used for dimensioning the objects. Though this specification, for simplicity and consistency, will refer to such dimensioning system **110** as a server, which typically performs operations of the present disclosure, those of ordinary skill in the art will appreciate that the disclosed dimensioning system **110** can also be implemented in various other computing systems like a computer stationed in a premises of an industry or a manufacturing plant, a mobile device of an operator responsible for performing various operations (e.g. quality check) of components/products manufactured in the industry or manufacturing plant, or a dedicated sensing device placed above/along with a conveyor belt carrying the manufactured components/products, or in any other environment in which dimensioning of the objects is required.

[0027] As can be seen from **Fig. 1**, the dimensioning system **110** is first trained during a training phase (left-hand side of **Fig. 1**) and then implemented during real-time operation (right-hand side of **Fig. 1**). In both the phases (training or real-time), the dimensioning system **110** may be in communication with one or more data sources (not shown in the figure) having reference images or real-time images. Additionally, the dimensioning system **110** may be in communication with other computing systems (not shown) or imaging devices via at least one network (not shown).

[0028] The network may comprise a data network such as, but not restricted to, the Internet, Local Area Network (LAN), Wide Area Network (WAN), Metropolitan Area Network (MAN), etc. In certain embodiments, the network may include a wireless network, such as, but not restricted to, a cellular network and may employ various technologies including Enhanced Data rates for Global Evolution (EDGE), General Packet Radio Service (GPRS), Global System for Mobile Communications (GSM), Internet protocol Multimedia Subsystem (IMS), Universal Mobile Telecommunications System (UMTS) etc. In one embodiment, the network may include or otherwise cover networks or subnetworks, each of which may include, for example, a wired or wireless data pathway.

[0029] In one embodiment, the one or more data sources may comprise a plurality of reference images which may be provided to the dimensioning system **110** for training purposes during the training phase. Whereas, in another embodiment, the one or more data sources may also comprise the real-time image provided to the dimensioning system **110** for dimensioning during the real-time operation.

[0030] Now, **Figure 1** is explained in conjunction with **Figure 2**, which is a detailed block diagram **200** of the dimensioning system **110** for dimensioning an object, in accordance with some embodiments of the present disclosure. According to an embodiment of the present disclosure, the dimensioning system **110** may comprise an imaging unit **202**, one or more interfaces **204**, at least one processor **206**, a memory **208**, a region identification model **112**, one or more auxiliary classifiers **114**, and various units **210**. The units **210** may comprise an implementing unit **212**, an applying unit **214**, a determining unit **216**, and various other units **218**. The various other units **218** may perform operations like receiving images from the data sources, training the region identification model and any other operations of the dimensioning system **110**. In an embodiment, the units **212-218** may be dedicated hardware units capable of executing one or more instructions stored in the memory **208** for performing various operations of the dimensioning system **110**. In another embodiment, the units **212-218** may be software modules stored in the memory **208** which may be executed by the at least one processor **206** for performing the operations of the dimensioning system **110**. In one embodiment, some or all of the functionalities of the units **212-218** may be performed by the at least one processor **206**.

[0031] The interfaces **204** may include a variety of software and hardware interfaces, for example, a web interface, a graphical user interface, an input device-output device (I/O) interface, a network interface and the like. The I/O interfaces may allow the dimensioning system **110** to interact with other computing systems directly or through other devices. The network interface may allow the dimensioning system **110** to interact with one or more data sources either directly or via a network.

[0032] The memory **208** may comprise various types of data. For example, the memory **208** may comprise ratio **116** obtained based on real dimensions and an image dimensions of a reference object. The memory **208** may further store one or more instructions executable by the at least one processor **206**. The memory **208** may be communicatively coupled to the at least one processor **206**. The memory **208** may include a Random-Access Memory (RAM) unit and/or a non-volatile memory unit such as a Read Only Memory (ROM), optical disc drive, magnetic disc drive, flash memory, Electrically Erasable Read Only Memory (EEPROM), a memory space on a server or cloud and so forth.

[0033] The at least one processor **206** may include, but not restricted to, a general-purpose processor, a Field Programmable Gate Array (FPGA), an Application Specific Integrated Circuit (ASIC), a Digital Signal Processor (DSP), microprocessors, microcomputers, micro-controllers, central processing units, state machines, logic circuitries, and/or any devices that manipulate signals based on operational instructions. A processor may also be implemented as a combination of computing devices, e.g., a combination of a DSP and a microprocessor, a plurality of microprocessors, one or more microprocessors in conjunction with a DSP core, or any other such configuration.

[0034] Now referring back to **Fig. 1**, it can be seen that the dimensioning system **110** is functioning in two phases – training and real-time. In the forthcoming paragraphs, both the phases of the dimensioning system **110** are explained by means of **Figures 1** and **2**.

Training phase

[0035] Initially, the dimensioning system **110** is trained using the plurality of reference images taken corresponding to a plurality of reference objects. For simplicity, only one reference object and its corresponding reference image is shown in **Fig. 1**. In one embodiment, the dimensioning system **110** may use the imaging unit **202** for capturing the reference images of the reference objects for training purposes. The imaging unit **202** may be an integral part of the dimensioning system **110** or external to the dimensioning system **110**. The imaging unit **202** further transmits the captured reference image to the at least one processor **206** for further processing. The imaging unit **202** may be an image capturing device like a camera capable of capturing the reference image once it comes in its field of view (FOV). However, according to an alternative embodiment, the at least one processor **206** may receive the reference image from one or more data sources like an external database or data storing device.

[0036] Once the reference image (of reference object) is received, the at least one processor **206** identifies a region of interest (ROI) in the reference image. The ROI is a particular region, within the image, which indicates the presence of the object. In next step, the at least one processor **206** determines a bounding box for the ROI in the image. According to an embodiment, the bounding box comprises a set of pixel coordinates corresponding to vertices of the bounding box of the reference object. However, the bounding box created in first instance is not precise. Stated another way, the bounding box created lacks accuracy and may not precisely brings the reference object under its purview, thereby leading to dimensioning error. To address this concern, the present disclosure implements a combination of region identification model (i.e. RPN model) **112** and the one or more auxiliary classifiers **114** to adjust the bounding box into a “new bounding box” corresponding to the ROI of the reference object. Since the new bounding box comprises a “new set of pixel coordinates” which are more precise, the corresponding length and width obtained using them are also precise. In next step, at least one processor **206** estimates one or more Euclidean distance based on the new set of pixel coordinates such that each Euclidean distance provides dimension between vertices of the new bounding box. The at least one processor **206** further transforms each Euclidean distance into dimension values (i.e. length and width) to determine the image dimensions of the reference object. Here, the image dimensions is nothing but an apparent dimension of the reference object. That is, the image dimensions/apparent dimensions is still not the exact dimension of the object.

[0037] However, since the dimensioning system **110** in the training phase, it leverages the advantage of using the real dimension of the reference object which is already known to it. Thus, the at least one processor **206** determines a “ratio **116**” between the real dimension and the image dimension (apparent) of the reference object. It may be noted that, the ratio **116** determined is more accurate i.e. closer to value 1 because of the image dimension was calculated based on the new set of pixel coordinates of the new bounding box, as described in the above paragraph. It may also be noted that the implementation the combination of region identification model (i.e. RPN model) **112** and the one or more auxiliary classifiers **114** is done twice – one during the training phase and another during the real-time operation. In the first instance, the combination helps in determining a more accurate ratio **116** which is further utilized while dimensioning the objects during the real-time. In second instance, the

combination helps in determining image dimensions (i.e. apparent dimension) more precisely for the real-time object, which will be now explained in following paragraphs.

Real-time operation

[0038] By the time the dimensioning system **110** is implemented in real-time, it gets trained in recognizing and measuring new objects. At one hand, the dimensioning system **110** now have better ratio **116** (obtained during the training phase) which is now utilized for dimensioning the objects in real-time. On the other hand, the dimensioning system **110** also have an advantage of the combination of RPN model **112** and the auxiliary classifiers **114** which is utilized for calculating image dimensions (apparent dimensions) of the real-time objects more accurately. The operation starts with detecting a real-time object when it comes in the field of view (FOV) of the imaging unit **202**. As described earlier, the imaging unit **202** (e.g. camera) may be associated with the dimensioning system **110** or may be an independent unit capable of capturing the images of the objects and transmitting them to the dimensioning system **110** for dimensioning. Whereas, in alternative scenario, the dimensioning system **110** will receive the real-time image from one or more data sources.

[0039] In either of the above scenarios, the at least one processor **206** first implements the region identification model **112** on the real-time image which is to be dimensioned. Based on the implementation, the at least one processor **206** identifies a region of interest (ROI) in the real-time image of the object. Stated earlier, the ROI is nothing but a particular region, within the image, which is indicates the presence of the object. Thereafter, the at least one processor **206** determines a bounding box for the ROI in the real-time image. The bounding generated is shown in bottom part of **Fig. 1**. It can be seen that the bounding box generated (shown as solid lines) comprises a first set of pixel coordinates corresponding to vertices of the bounding box and a first overlap value. According to embodiments of present disclosure, the first overlap value is nothing but an Intersection over Union (IoU) value which is obtained by calculating a ratio of an area of an intersection between a predicted bounding box and a ground-truth bounding box to an area of a union between the predicted bounding box and the ground-truth bounding box. In other words, the first overlap value (IoU) indicates a degree of overlap of the bounding box generated upon the object which is to be dimensioned. The first pixel coordinates and the first overlap value are shown in below table 1.

Vertices of bounding box→	TL ₁	TR ₁	BR ₁	BL ₁
First Pixel coordinates→	(143, 223)	(196, 225)	(195, 103)	(142, 102)
First Overlap Value→	59.8			

Table 1 – First set of pixel coordinates and first overlap value of the bounding box (before adjustment)

[0040] It can be understood that the bounding box generated above is solely based on the implementation of the region identification model **112** i.e. RPN model. Therefore, the bounding box generated have some error or less precision, which is identified as a technical problem sought to be

addressed by the dimensioning technique disclosed in the present disclosure. In other words, the bounding box generated in first stage may not properly localize the object which is to be dimensioned. Such localization error may lead to dimensioning error as ultimately length and width of the bounding box generated will be considered for dimensioning the object.

[0041] To address this technical challenge, the disclosed dimensioning technique uses the combination of region identification model (i.e. RPN model) 112 and the one or more auxiliary classifiers 114. As described earlier, this combination is used in two instances – first during the training phase and second during the real-time operation. And in each instances the combinations achieves a technical advantage i.e. more precisely determining image dimension (apparent) of the reference object and arriving at an accurate ratio 116 (during training phase as described earlier) and again more precisely determining the image dimension (apparent) during the real-time operation, which is currently being discussed.

[0042] Thus, when the one or more auxiliary classifiers 114 are applied upon the region identification model 112, the model 112 adjust the bounding box to a new bounding box (shown as dotted lines) such that new bounding box comprises a second set of pixel coordinates and a second overlap value. The adjustment is clearly shown in Fig. 1, in which, the new bounding box “shrinks” from all the sides to localize the object more accurately. Like the first overlap value, the second overlap value is also an Intersection over Union (IoU) value which is obtained by calculating a ratio of an area of an intersection between a predicted bounding box and a ground-truth bounding box to an area of a union between the predicted bounding box and the ground-truth bounding box. In other words, the second overlap value (IoU) indicates a degree of overlap of the new bounding box generated upon the object which is to be dimensioned. The second pixel coordinates and the second overlap value are shown in below table 2.

Vertices of bounding box →	TL ₂	TR ₂	BR ₂	BL ₂
Second Pixel coordinates →	(140, 220)	(193, 220)	(193, 100)	(140, 100)
Second Overlap Value →	74.6			

Table 2 – Second set of pixel coordinates and second overlap value of the new bounding box (before adjustment)

[0043] Now on comparing table 1 and table 2, it can be observed that the second pixel coordinates are more precise than the first pixel coordinates. In other words, the length and width of the new bounding box generated will be more precise. Also, the second overlap value i.e. “74.6” is greater than the first overlap value i.e. “59.8”. This clearly indicates that the new bounding box have a higher degree of overlap compared to the bounding box before adjustment. Though in this example as shown in Fig. 1, the new bounding box is shown to be shrinking from all the sides, however those of ordinary skill in the art will appreciate that the new bounding box may be formed in different transformation like expanding from all the sides or a combination of shrinking and expanding from one or more sides of

the new bounding box depending upon the correction being done for correctly localizing the object in the image.

[0044] In next step, the at least one processor **206** determines an image dimensions of the real-time object based on the second set of pixel coordinates of the new bounding box or second bounding box generated after the adjustment. It may be understood that the image dimensions calculated above is an apparent dimension, and therefore, it is still not the actual dimension of the real-time object. However, the image dimensions calculated for the real-time object is more precise compared to the one which would have been calculated using the existing image dimensioning technique i.e. simply implementing the RPN model **112** without applying the one or more classifiers **114**. Now with the more precise image dimensions being calculated (during the real-time operation) and with the already stored ratio **116** (determined during the training phase), the dimensioning system **110** comes in much better position to accurately determine the real dimension of the real-time object. Thus, the at least one processor **206** determines the real dimensions of the real-time object using the image dimensions of the real-time object and the ratio **116**.

[0045] This way, the present disclosure provides faster and accurate dimensioning of the objects in the real-time. By faster, it means that the dimensions of the objects can be determined as soon as they come in the field of view (FOV) of the imaging unit **202** associated with the dimensioning system **110**. By accurate, it means that by implementing the auxiliary classifiers **114** into the region identification model (e.g. RPN model) **112**, the disclosed dimensioning technique accurately determines the dimensions of the object. Such implementation further provides a technical improvement in the existing image dimensioning technique. Also, the disclosed dimensioning technique can be implementing in an existing hardware setup/ environment in the manufacturing plants or industries, thereby providing an efficient and cost effective technique.

[0046] Referring now to **Figure 3** which illustrates a flow diagram for dimensioning an object, in accordance with some embodiments of the present disclosure. The method **300** is merely provided for exemplary purposes, and embodiments are intended to include or otherwise cover any methods or procedures for evaluating one or more functionalities of an application.

[0047] The method **300** may include, at block **302**, implementing a region identification model **112** on an image of the object to be dimensioned. The implementing comprises identifying a region of interest (ROI) in the image of the object and further determining a bounding box for the ROI in the image. The bounding box comprises a first set of pixel coordinates corresponding to vertices of the bounding box and a first overlap value indicating a degree of overlap of the bounding box upon the object. The operations of block **302** may be performed by the at least one processor **206** or an implementing unit **212** of **Figure 2**.

[0048] At block **304**, the method **300** may include applying one or more auxiliary classifiers **114** upon the region identification model **112** in order to enable the region identification model **112** to adjust the bounding box to a new bounding box such that new bounding box comprises a second set of pixel

coordinates and a second overlap value. The second overlap value is greater than the first overlap value, thereby indicating that the new bounding box having higher degree of overlap compared to the bounding box before adjustment. The operations of block **304** may be performed by the at least one processor **206** or an applying unit **214** of **Figure 2**.

[**0049**] At block **306**, the method **300** may include determining an image dimensions of the object based on the second set of pixel coordinates of the second bounding box. The operations of block **306** may be performed by the at least one processor **206** or a determining unit **216** of **Figure 2**.

[**0050**] At block **308**, the method **300** may include determining a real dimensions of the object based on the image dimensions of the object, and a ratio **116** obtained between a real dimensions and an image dimensions of a reference object. The operations of block **308** may be performed by the at least one processor **206** or the determining unit **216** of **Figure 2**.

[**0051**] Referring now to **Figure 4** which illustrates a flow diagram for determining a ratio, in accordance with some embodiments of the present disclosure. The method **400** is merely provided for exemplary purposes, and embodiments are intended to include or otherwise cover any methods or procedures for evaluating one or more functionalities of an application.

[**0052**] The method **400** may include, at block **402**, receiving an image, of the reference object, captured by an imaging unit **202** when the object is placed in a field of view (fov) of the imaging unit **202**. Whereas, the real dimensions of the reference object is known. The operations of block **402** may be performed by the at least one processor **206** of **Figure 2**.

[**0053**] At block **404**, the method **400** may include identifying a region of interest (ROI) in the image of the reference object. The operations of block **404** may be performed by the at least one processor **206** of **Figure 2**.

[**0054**] At block **406**, the method **400** may include determining a bounding box for the ROI in the image. The bounding box comprises a set of pixel coordinates corresponding to vertices of the bounding box of the reference object. The operations of block **406** may be performed by the at least one processor **206** of **Figure 2**.

[**0055**] At block **408**, the method **400** may include implementing the region identification model **112** along with the one or more auxiliary classifiers **114** to adjust the bounding box into a new bounding box corresponding to the ROI of the reference object. The new bounding box comprises a new set of pixel coordinates. The operations of block **408** may be performed by the at least one processor **206** of **Figure 2**.

[**0056**] At block **410**, the method **400** may include determining an image dimensions of the reference object based on the new set of pixel coordinates of the new bounding box. The operations of block **410** may be performed by the at least one processor **206** of **Figure 2**.

[0057] At block **412**, the method **400** may include determining the ratio **116** between the real dimensions and the image dimensions of the reference object such that the ratio **116** is utilizable to dimension objects other than the reference object. The operations of block **412** may be performed by the at least one processor **206** of **Figure 2**.

[0058] The above methods **300 and 400** may be described in the general context of computer executable instructions. Generally, computer executable instructions can include routines, programs, objects, components, data structures, procedures, modules, and functions, which perform specific functions or implement specific abstract data types.

[0059] The order in which the various operations of the method are described is not intended to be construed as a limitation, and any number of the described method blocks can be combined in any order to implement the method. Additionally, individual blocks may be deleted from the methods without departing from the spirit and scope of the subject matter described herein. Furthermore, the methods can be implemented in any suitable hardware, software, firmware, or combination thereof.

[0060] Furthermore, one or more computer-readable storage media may be utilized in implementing embodiments consistent with the present disclosure. A computer-readable storage medium refers to any type of physical memory on which information or data readable by a processor may be stored. Thus, a computer-readable storage medium may store instructions for execution by one or more processors, including instructions for causing the processor(s) to perform steps or stages consistent with the embodiments described herein. The term “computer-readable medium” should be understood to include tangible items and exclude carrier waves and transient signals, i.e., non-transitory. Examples include Random Access Memory (RAM), Read-Only Memory (ROM), volatile memory, nonvolatile memory, hard drives, Compact Disc (CD) ROMs, Digital Video Disc (DVDs), flash drives, disks, and any other known physical storage media.

[0061] Certain aspects may comprise a computer program product for performing the operations presented herein. For example, such a computer program product may comprise a computer readable media having instructions stored (and/or encoded) thereon, the instructions being executable by one or more processors to perform the operations described herein. For certain aspects, the computer program product may include packaging material.

[0062] Various components, modules, or units are described in this disclosure to emphasize functional aspects of devices configured to perform the disclosed techniques, but do not necessarily require realization by different hardware units. Rather, as described above, various units may be combined in a hardware unit or provided by a collection of interoperative hardware units, including one or more processors as described above, in conjunction with suitable software and/or firmware.

[0063] Finally, the language used in the specification has been principally selected for readability and instructional purposes, and it may not have been selected to delineate or circumscribe the inventive

subject matter. It is therefore intended that the scope of the invention be limited not by this detailed description, but rather by any claims that issue on an application based here on. Accordingly, the embodiments of the present invention are intended to be illustrative, but not limiting, of the scope of the invention, which is set forth in the appended claims.

CLAIMS

We Claim

1. A method for dimensioning an object, the method comprising:
 - implementing (302) a region identification model (112) on an image of the object to be dimensioned, wherein implementing the region identification model (112) comprises:
 - identifying a region of interest (ROI) in the image of the object;
 - determining a bounding box for the ROI in the image, wherein the bounding box comprises a first set of pixel coordinates corresponding to vertices of the bounding box and a first overlap value indicating a degree of overlap of the bounding box upon the object;
 - applying (304) one or more auxiliary classifiers (114) upon the region identification model (112) in order to enable the region identification model (112) to:
 - adjust the bounding box to a new bounding box such that the new bounding box comprises a second set of pixel coordinates and a second overlap value, wherein the second overlap value is greater than the first overlap value, thereby indicating that the new bounding box having higher degree of overlap compared to the bounding box before adjustment;
 - determining (306) an image dimensions of the object based on the second set of pixel coordinates of the second bounding box; and
 - determining (308) a real dimensions of the object based on the image dimensions of the object, and a ratio (116) obtained between a real dimensions and an image dimensions of a reference object.
2. The method as claimed in claim 1, wherein the ratio (116) is determined by:
 - receiving (402) an image, of the reference object, captured by an imaging unit (202) when the object is placed in a field of view (fov) of the imaging unit (202), wherein the real dimensions of the reference object is known;
 - identifying (404) a region of interest (ROI) in the image of the reference object;
 - determining (406) a bounding box for the ROI in the image, wherein the bounding box comprises a set of pixel coordinates corresponding to vertices of the bounding box of the reference object;
 - implementing (408) the region identification model (112) along with the one or more auxiliary classifiers (114) to adjust the bounding box into a new bounding box corresponding to the ROI of the reference object, wherein the new bounding box comprises a new set of pixel coordinates;
 - determining (410) an image dimensions of the reference object based on the new set of pixel coordinates of the new bounding box; and
 - determining (412) the ratio (116) between the real dimensions and the image dimensions of the reference object, wherein the ratio is utilizable to dimension objects other than the reference object.
3. The method as claimed in claim 2, wherein determining the image dimensions of the reference object further comprises:

estimating one or more Euclidean distance based on the new set of pixel coordinates such that each Euclidean distance provides dimension between vertices of the new bounding box; and

transforming each Euclidean distance into dimension values to determine the image dimensions of the reference object.

4. The method as claimed in claim 1, wherein the one or more auxiliary classifiers (114) are convolution neural network classifiers.
5. The method as claimed in claim 1, wherein the first overlap value and the second overlap value are obtained by calculating a ratio of an area of an intersection between a predicted bounding box and a ground-truth bounding box to an area of a union between the predicted bounding box and the ground-truth bounding box.
6. A dimensioning system (110) for dimensioning an object, wherein the dimensioning system (110) comprises:
 - a memory (208); and
 - at least one processor (206) operatively coupled to the memory (208) and configured to:
 - implementing a region identification model (112) on an image of the object to be dimensioned, wherein implementing the region identification model (112) comprises:
 - identifying a region of interest (ROI) in the image of the object;
 - determining a bounding box for the ROI in the image, wherein the bounding box comprises a first set of pixel coordinates corresponding to vertices of the bounding box and a first overlap value indicating a degree of overlap of the bounding box upon the object;
 - apply one or more auxiliary classifiers (114) upon the region identification model (112) in order to enable the region identification model (112) to:
 - adjust the bounding box to a new bounding box such that the new bounding box comprises a second set of pixel coordinates and a second overlap value, wherein the second overlap value is greater than the first overlap value, thereby indicating that the new bounding box having higher degree of overlap compared to the bounding box before adjustment;
 - determine an image dimensions of the object based on the second set of pixel coordinates of the second bounding box; and
 - determine a real dimensions of the object based on the image dimensions of the object, and a ratio (116) obtained between a real dimensions and an image dimensions of a reference object.
7. The dimensioning system (110) as claimed in claim 6, further comprises an imaging unit (202) coupled with the at least one processor (206), wherein the at least one processor (206), for determining the ratio (116), is configured to:

receiving an image, of the reference object, captured by the imaging unit when the object is placed in a field of view (fov) of the imaging unit, wherein the real dimensions of the reference object is known;

identify a region of interest (ROI) in the image of the reference object;

determine a bounding box for the ROI in the image, wherein the bounding box comprises a set of pixel coordinates corresponding to vertices of the bounding box of the reference object;

implement the region identification model (112) along with the one or more auxiliary classifiers (114) to adjust the bounding box into a new bounding box corresponding to the ROI of the reference object, wherein the new bounding box comprises a new set of pixel coordinates;

determine an image dimensions of the reference object based on the new set of pixel coordinates of the new bounding box; and

determine the ratio (116) between the real dimensions and the image dimensions of the reference object, wherein the ratio (116) is utilizable to dimension objects other than the reference object.

8. The dimensioning system (110) as claimed in claim 7, wherein the at least one processor (206), to determine the image dimensions of the reference object, is further configured to:
 - estimate one or more Euclidean distance based on the new set of pixel coordinates such that each Euclidean distance provides dimension between vertices of the new bounding box; and
 - transform each Euclidean distance into dimension values to determine the image dimensions of the reference object.
9. The dimensioning system (110) as claimed in claim 6, wherein the one or more auxiliary classifiers (114) are convolution neural network classifiers.
10. The dimensioning system (110) as claimed in claim 6, wherein the first overlap value and the second overlap value are obtained by calculating a ratio of an area of an intersection between a predicted bounding box and a ground-truth bounding box to an area of a union between the predicted bounding box and the ground-truth bounding box.

Dated this 11th Day of March 2022

Robin Koshy Varghese (INPA No: 3705)

Head, IPR Dept.

L&T Technology Services Ltd.

DLF 3rd Block, 2nd Floor,

Manapakkam, Chennai - 600089.

ABSTRACT

“METHOD AND SYSTEM FOR DIMENSIONING AN OBJECT”

The present disclosure describes technique for dimensioning an object. The technique includes implementing region identification model on image of the object to identify ROI in the image and a bounding box for the ROI. The bounding box comprises first set of pixel coordinates corresponding to vertices of the bounding box and first overlap value indicating degree of overlap of the bounding box upon the object. The technique includes applying auxiliary classifiers upon the region identification model to adjust the bounding box to new bounding box comprising second set of pixel coordinates and second overlap value greater than the first overlap value. The technique includes determining image dimensions of the object based on the second set of pixel coordinates of the second bounding box. The real dimensions of the object is determined based on the image dimensions of the object, and ratio obtained between real and image dimensions of reference object.

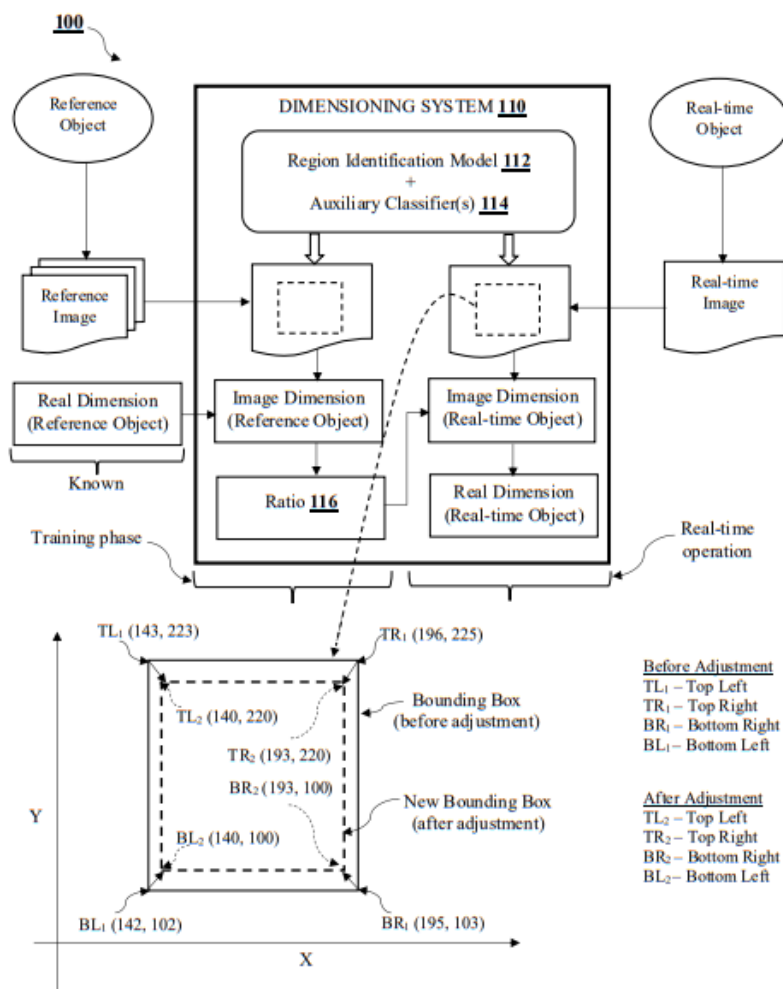


Figure 1

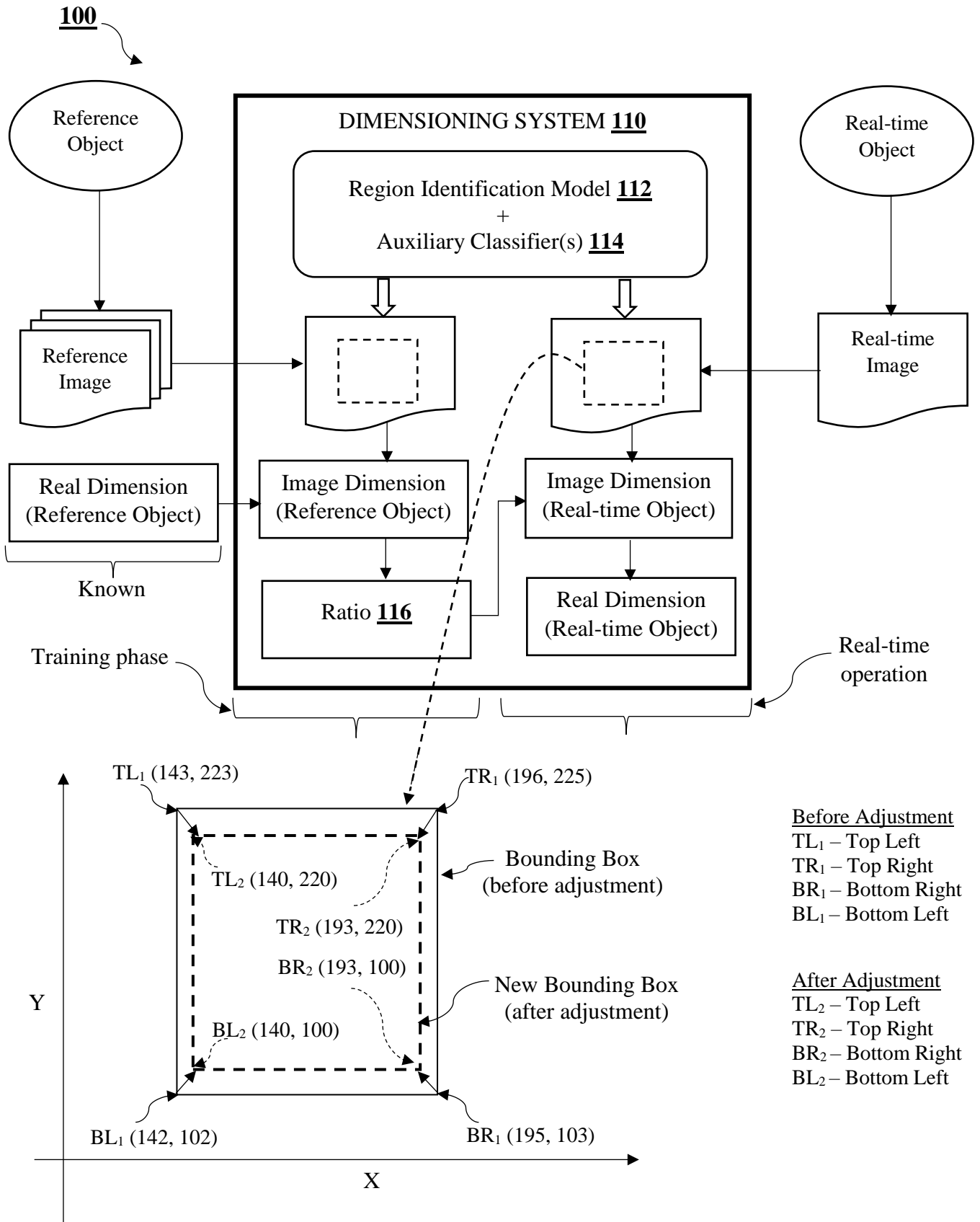


Figure 1

Robin Koshy Varghese (INPA No: 3705)
Head, IPR Dept.,
L&T Technology Services Limited,
DLF 3rd Block, 2nd Floor,
Manapakkam, Chennai - 600089.

200

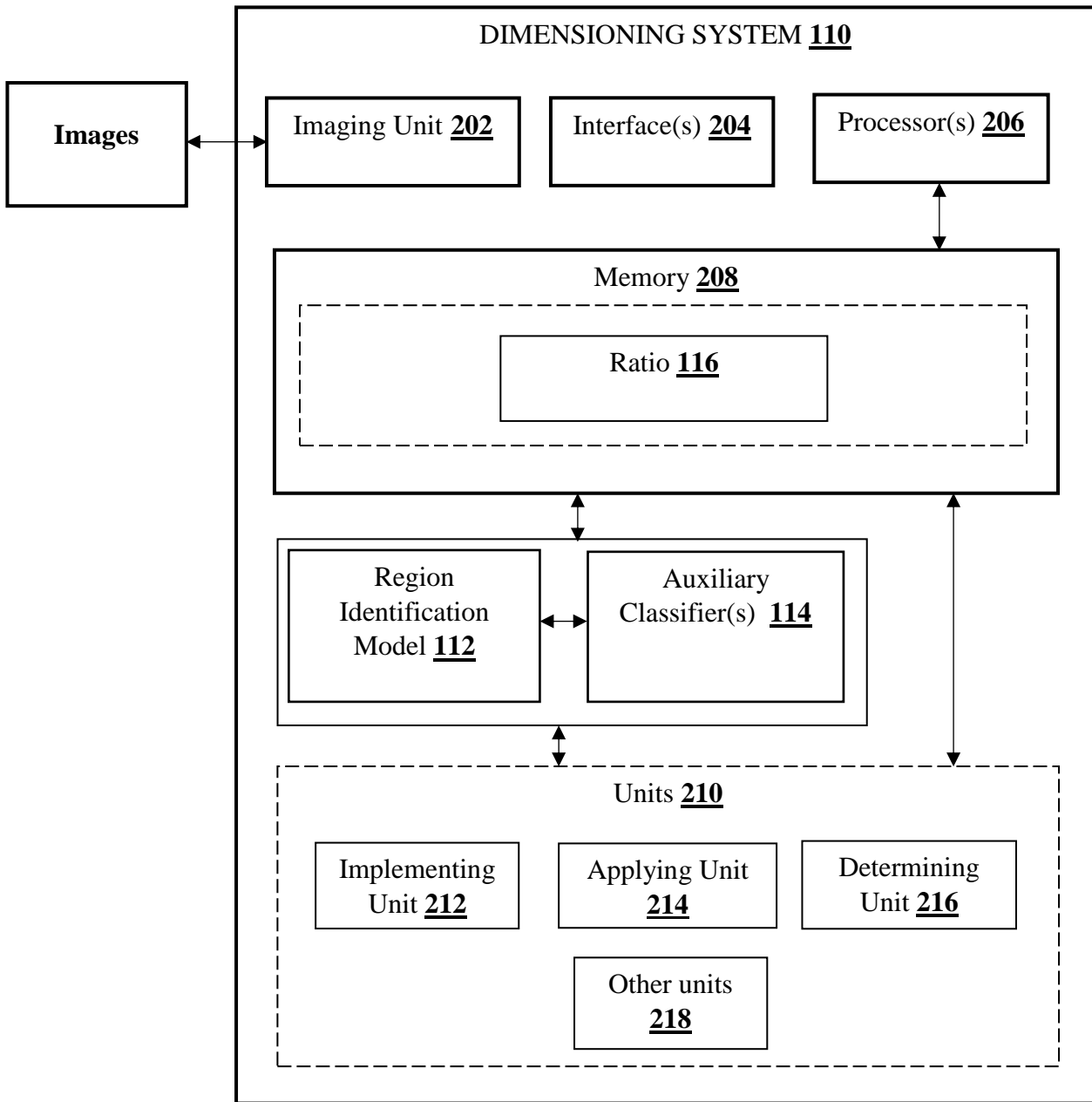
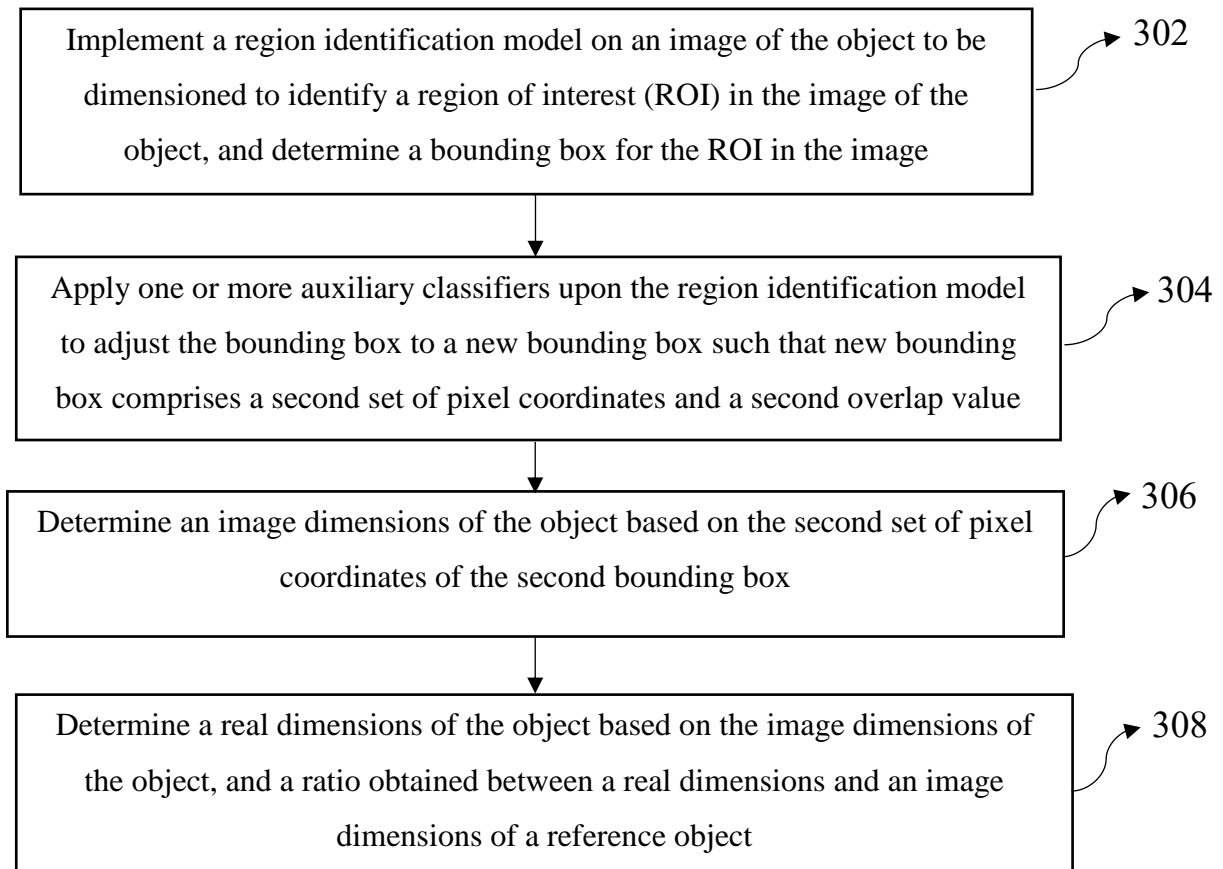
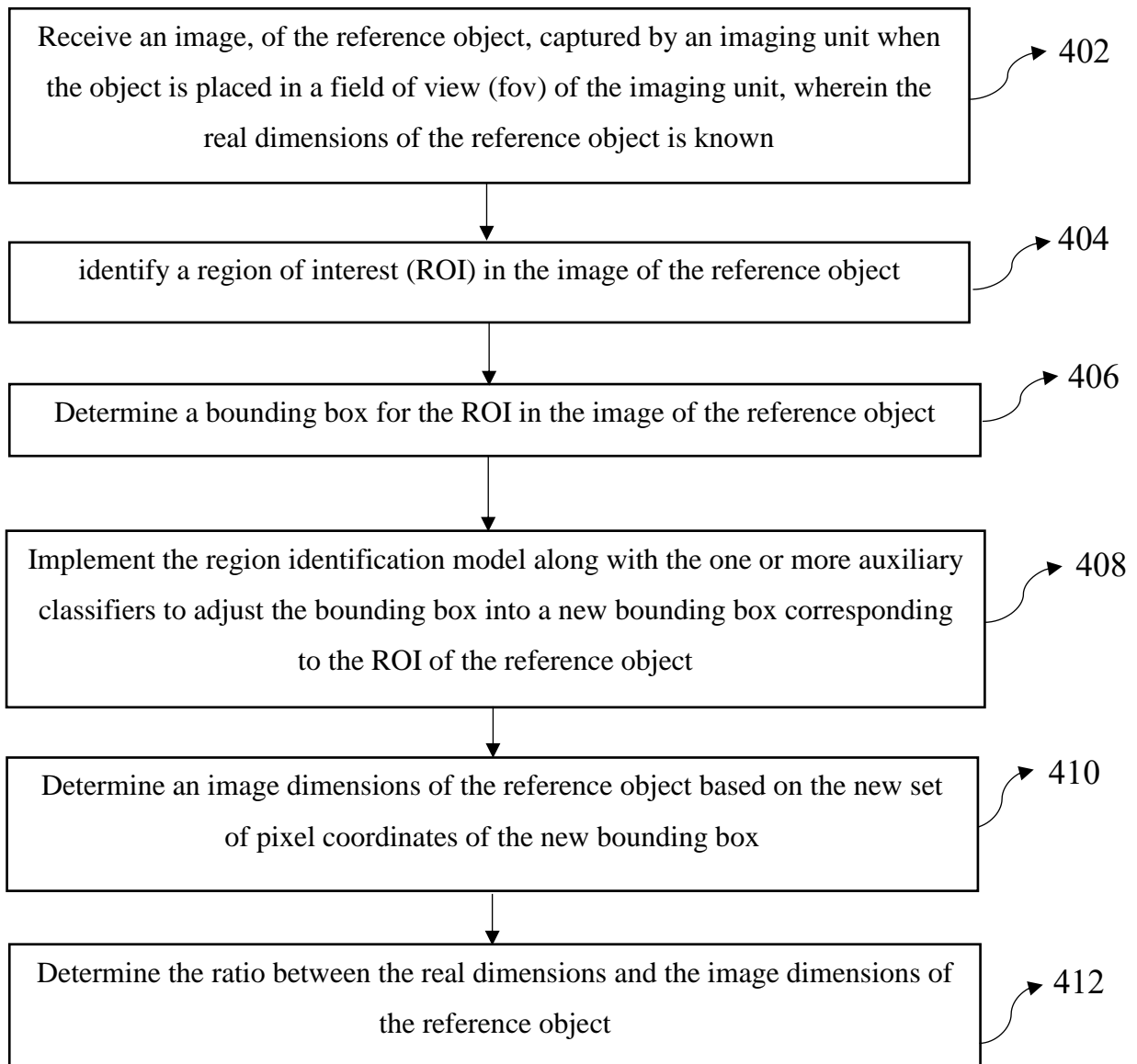


Figure 2

Robin Koshy Varghese (INPA No: 3705)
Head, IPR Dept.,
L&T Technology Services Limited,
DLF 3rd Block, 2nd Floor,
Manapakkam, Chennai - 600089.

300**Figure 3**

**Robin Koshy Varghese (INPA No: 3705)
Head, IPR Dept.,
L&T Technology Services Limited,
DLF 3rd Block, 2nd Floor,
Manapakkam, Chennai - 600089.**

400**Figure 4**

Robin Koshy Varghese (INPA No: 3705)
Head, IPR Dept.,
L&T Technology Services Limited,
DLF 3rd Block, 2nd Floor,
Manapakkam, Chennai - 600089.