

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization

International Bureau

(43) International Publication Date
03 May 2018 (03.05.2018)



(10) International Publication Number
WO 2018/078613 A1

(51) International Patent Classification:

B65B 57/14 (2006.01) *G06K 9/46* (2006.01)

Published:

— with international search report (Art. 21(3))

(21) International Application Number:

PCT/IL2017/050990

(22) International Filing Date:

04 September 2017 (04.09.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

248641 31 October 2016 (31.10.2016) IL

(71) Applicants: **D.I.R. TECHNOLOGIES (DETECTION IR) LTD.** [IL/IL]; Matam Towers 3, POB 15129, 3190501 Haifa (IL). **RAFAEL ADVANCED DEFENSE SYSTEMS LTD.** [IL/IL]; P.O.B. 2250, 3102102 Haifa (IL).

(72) Inventors: **PINSKY, Ephraim**; 44 Ya'ara Street, 3652072 Kiryat Tivon (IL). **PELES, David**; 31/28 Oren Street, 3473502 Haifa (IL). **MARKOWITZ, Zvi**; 12/20 Medinat Ha-Yehudim Street, 4676612 Herzliya (IL). **IMMER, Efrat**; 16/9 Ben Sira Street, 6291625 Tel Aviv (IL).

(74) Agent: **LUZZATTO, Kfir** et al.; Luzzatto & Luzzatto, P.O. Box 5352, 8415202 Beer Sheva (IL).

(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, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— of inventorship (Rule 4.17(iv))

(54) Title: SYSTEM AND METHOD FOR AUTOMATIC INSPECTION AND CLASSIFICATION OF DISCRETE ITEMS

(57) Abstract: A method for inspecting and classifying discrete items comprises capturing the images of a free-falling item using at least one imaging device and analyzing the captured images using a neural network which has been trained using images and/or simulated images of regular products, defected products and Mix-Ups.



WO 2018/078613 A1

SYSTEM AND METHOD FOR AUTOMATIC INSPECTION AND
CLASSIFICATION OF DISCRETE ITEMS

Field of the Invention

The present invention relates to a system and to a method for the automatic inspection and classification of discrete items. More particularly, the invention relates to inspection and classification apparatus and method that are useful to identify discrete items possessing a specific characteristic in the course of the packaging and/or manufacturing thereof.

Background of the Invention

Many different items are handled in industry by automatic apparatus, for a variety of purposes. For instance, packaging of a give number of screws in a package must be monitored by counting the number of screws going into it. Seeds used in agriculture, on the other hand, must be checked also for viability so that a chipped or broken seed is not planted, resulting in lost growth. In automated packaging facilities and factories, the number of items processed per second is typically very high, to achieve economic efficiency, and therefore the problem of inspecting them is essentially universal.

A particularly sensitive field related to the above is that of pharmaceutical and nutraceutical products, because in addition to the need to assure that the correct number of items (pills, tablets, lozenges, capsules, etc.) is packaged and that the item is intact (i.e., is not broken, deformed, discoloured or chipped – which is defined as "Major defect"), this field is particularly sensitive to the problem of Mix-Up – i.e., an event in which a foreign item or object becomes mixed up with the correct ones (Defined as "Critical problem"). While this is a universal problem (for instance, you don't want a bolt to get mixed with a set of screws), it is particularly problematic in the field of pharmaceuticals because it may lead to a patient taking the wrong medicine, with the resulting potential health hazards.

The art has so far failed to provide an efficient method and system that not only performs all other required tasks, such as counting and detecting defects in an item, but efficiently prevents Mix-Ups with a high level of success. It is therefore clear that it would be highly desirable to provide such a system, which may ensure a high degree of confidence in avoiding Mix-Ups and defects.

It is a purpose of the present invention to provide a method and system, which are effective in preventing Mix-Ups and major defects in the packaging phase carried out by high capacity filling machines.

It is another purpose of the invention to provide such a method and system that may prevent the unwanted packaging of defective products in bottles (or other containers) of solid dosage products, such as bottles of tablets or capsules.

It is a further object of the invention to provide a method and system suitable to distinguish between at least two different levels of severity in order to trigger a full stop of the filling machine or rejection of products or a bottle of products or other container.

It is yet another purpose of the invention to provide a method and system suitable to count the number of products which enter each bottle while it is being filled, thereby to control the bottle filling process or to reject bottles with wrong count of products.

All the above and other characteristics and advantages of the invention will become apparent as the description proceeds.

Summary of the Invention

The invention relates to a method for inspecting and classifying discrete items, comprising capturing the images of a free-falling item using at least one imaging device and analyzing the captured images using a neural

network which has been trained using images and/or simulated images of regular products, defected products and Mix-Ups. The term "simulated", as used herein, includes all types of non-real life images, such as "augmented" images or composite images.

According to one embodiment of the invention the item's free fall occurs from a filling machine.

In another embodiment of the invention images from two or more imaging devices are fed to the classifier.

The discrete item can be of any kind. One of the cases in which the invention is particularly useful is when the discrete article is a pharmaceutical or nutraceutical product.

Items classification may vary according to the actual nature of the item. However, for many practical applications, such as in many cases of filling operations of pharmaceutical or nutraceutical products, three classes are particularly useful, i.e., items that are "Good", have a Major Defect, or have a Critical Problem. In one embodiment of the invention if a Critical Problem is detected, the operation of the filling machine is stopped, whereas if a Major defect is detected, one or more packaging that have been filled prior to such detection are removed from the production line.

In another aspect the invention is directed to a system for inspecting and classifying discrete items, comprising at least one imaging device suitable to capture the images of a free-falling item, and analysis apparatus suitable to analyze the images captured by said imaging device to assign a class to it according to a preset classification. In one embodiment of the invention one or more image device(s) are provided, which are suitable to capture the free-falling article before its free fall begins.

As said, the system of the invention comprises an analysis apparatus, which in one embodiment of the invention comprises a Decision Unit (DU) provided with a Processing Unit, which Decision Unit operates a neural network.

The system of the invention is provided with a database, which may include data generated by simulating a change in properties of items in the captured images. said change in properties may comprise a variation in size and/or color and/or shape and/or background of an item, generated from a real picture thereof. Illustrative items that can be processed by the system of the invention include tablets, pills and capsules, including chewy medicines.

In one embodiment of the invention the system comprises rotation-imparting elements suitable to cause the item to rotate during its fall, thereby permitting an extended view and inspection of the surface thereof. The invention also encompasses a method for training a classifier for classifying images or sequences of images of items comprising the steps of:

a) capturing a sequence of images which contain only a single type of either good or defected items or Mix-Ups; In this respect reference is also made herein to non-good classes recordings i.e. defected products or Mix-ups

b) segmenting said images using foreground-background segmentation;

c) generating additional images using one or more of the following options, alone or combined:

1) implanting segmented blobs of items from images of a specific type of items (either defected or Mix-Ups including simulated Mix-Ups) into images of "Good" items, in order to create images which contain multiple types of items;

2) simulating changes of lighting of the images;

3) simulating changes in the background of the items, including replacing the background with a constant background or replacing the background to match a background of a different machine or randomly change to the background;

4) simulating several levels of changes to the Good items shape, size and/or color including minor changes which are in the allowed tolerance of the "Good" items, and larger changes which exceed said allowed tolerance, including changes that slightly exceed said tolerance;

5) Simulating changes to Mix-Ups size and/or color and/or shape.

5) simulating minor shifts of the cameras;

6) generating images of rendered 3D models of items;

7) implanting rendered 3D models of items in captured images; and

d) using said captured and generated images to train a classifier for classifying images or sequences of images.

According to one embodiment of the invention a further step can be added, by filtering the training database by removing from the training database (DB) images of the non-good classes recordings in which a defect is not visible or no article is present in the image and also removing images from the good class record which might contain a defected product by one or more of the following:

1) using a foreground-background segmentation to analyze the presence, area and location of items in the images and removing from the training DB images of the non-good records in which no

product is presence in the image or only a small portion of the product overlaps the analysis region of interest (ROI) i.e. the area within the image which is analyzed by the classifier.

2) training a robust classifier to distinguish between images from the good record and images from the defected records and removing from the training DB images which got low score for the class which matches the record they were taken from.

3) same as in 2), but the DB is split into two sets and a separate classifier is trained on each of these sets and is activated on the second set in order to get a classification score for each of the classes.

4) same as in 2) or in 3), but a human observer is asked to confirm the removal of an image from the training DB.

All the above and other characteristics and advantages of the invention will be further understood by reference to the following description and appended figures.

Brief Description of the Drawings

In the drawings:

Fig. 1 illustrates a critical problem detected according to the invention, in which a Mix-Up causes a stop of the packaging/filling process;

Fig. 2 shows for comparison an image that contains no defects i.e., an image of a “Good” product;

Fig. 3 shows the presence of a hair on a tablet. Though a hair is a foreign object, some manufacturers might consider it as a major defect and not a critical problem, which will result in the discarding of the packaging into which it has fell;

Fig. 4 shows a major defect, in which a broken tablet is detected;

Fig. 5 shows a packaging apparatus—in this illustrative example, a high capacity bottle-filling machine—with capsules about to begin their fall toward their package;

Fig. 6 schematically shows the positioning of cameras according to one embodiment of the invention;

Fig. 7 is a schematic representation of a working system according to one embodiment of the invention;

Fig. 8 shows a Mix-Up artificially created by implanting an image of a product into an image that contains a different product, using image processing techniques; and

Fig. 9 shows changes in size and color made to a real-life image using image-processing techniques. These changes exceed the allowed tolerance of good products appearance and changes the label of the image from Good to Mix-up. This technique allows teaching the system to identify variations in shape and size and also increases the variety of the training examples. The same technique can be used to generate small

appearance changes which are within the allowed tolerance of good products and therefore do not change the label of the image. This technique may also be used to generate examples that slightly exceed the allowed tolerance of Good products and by that, generating "hard" training examples that help to improve the sensitivity of the resulting classifier.

Detailed Description of the Invention

While the invention is not limited to any particular filling apparatus, product or package, for the sake of brevity and clarity the invention will be illustrated hereinafter with reference to a high capacity bottle filling machine for pharmaceutical products, it being understood that the invention is not limited to such system in any way and can be applied to any other relevant system, with the appropriate adjustments.

Generally speaking, the method and system of the invention employ imaging devices (such as digital video cameras) which capture images of the products, a computer and computer vision algorithms that analyse the images and classify them as either: Good products, Critical problem or Major defect. While those are the common classes and, therefore, will be referred to for the purpose of illustration hereinafter, the invention may deal with other types of problems, not discussed herein for the sake of

brevity. As will be apparent to the skilled person the invention is not limited to any specific classification.

Two main types of high capacity filling machines – Slat machine and Electronic machines – are used in the industry. The invention will be illustrated with reference to an Electronic machine, but of course it applies, *mutatis mutandis*, to Slat machines as well.

In the illustrative embodiment illustrated below two fast and synchronized cameras are used to obtain a stereo image of the products. However, in other embodiments of the invention more cameras or a single camera may be used as well. The advantage of using stereo images is mainly that it allows capturing the product from more view angles, which allows inspection of a larger portion of the product's surface. By using a stereo image, 3D information may be used by the classification algorithm to classify the products. In this illustrative embodiment the cameras were positioned so that they will capture a portion of the end of the tray as well as the area of the free fall of the products from the end of the tray into the counting head (all of which will be described in detail with reference to Figs. 5 and 6). However, since the product's orientation typically changes during free fall, even using a single camera allows seeing it from different angles.

The classification algorithms may rely both on images of the products before it fell from the end of the tray and on images taken while the product was in free fall. Combining the information from several images allows inspecting more viewing angles of the product.

Referring now to Fig. 1, two items are seen in free fall, after they left edge 51 of tray 50 (Fig. 5). In this figure numeral 10 is the tablet that it is intended to fill, viewed from a different angle as 10a, while 11 is a capsule that got mixed up with the stream of tablets 11 (also shown from a different angle at 11a), triggering a “critical problem” signal from the system, which normally stops the operation of the filling machine. Critical problems must typically be thoroughly investigated, so in this case (pending other rules and instructions) it is not sufficient to remove the package into which the mixed-up capsule has fallen.

In contrast, Fig. 2 shows a process in which no problem is detected and all three tablets, 21, 22 and 23 (again, viewed from a different angle as 21a, 22a and 23a, respectively), appear to be uniform and without defects.

There are several advantages in inspecting the product while it falls:

1. The inspection is done as close as possible to the last stage of the packaging in order to minimize the chance that a problem would be formed after the inspection point. The free fall of the products into

the counting head or into the bottles is through a better inspection point than inspecting the product only on the tray.

2. The free fall of the products causes changes of the product angle and allows inspection of a larger portion of the product's surface.
3. The products are more separated from one another in the free fall stage, allowing less occlusions and better inspection as a result.
4. The product velocity at the beginning of the free falling stage is low, enabling sharp images capturing free of motion blur.
5. When inspecting in the free fall stage, it is possible to infer into which bottle the product has fallen. In contrast, detection of defect on the tray may not be easily matched to a specific bottle, requiring the system to reject several bottles or to stop the filling process.
6. It is possible to integrate a product rejection system in the free fall stage in order to reject a small number of products (instead of rejecting a bottle) with minimal delay to the filling process and minimal loss of products.

Fig. 3 illustrates a different kind of error. It can be easily seen that tablet 31 has a hair, 32, sticking to it. It can also be seen from this figure why it is advantageous to use a stereoscopic image and two cameras, since in the left-hand side, the hair is more clearly seen at 32a.

Fig. 4 illustrates a major defect, where tablet 41 is broken in half.

Fig. 5 shows a plurality of trays 50 carrying capsules that are driven (in this case by vibration feeding) to edge 51, from which they will fall into void 52 and, eventually, into a packaging, the top of which is shown at numeral 53.

Fig. 6 shows the positioning of two couples of cameras, 61-61a and 62-62a, pointing to two separate trays. The details of the system given in the figure (distances, L, and field of view, FOV) are for the illustrative example and are not meant to limit the invention in any way. Of course, employing different cameras with different characteristics will result in a different setup.

Fig. 7 is a scheme that illustrates the setup of a system according to a particular embodiment of the invention. In this embodiment two cameras (Camera 1 and Camera 2) are synchronized using synchronization unit SU that coordinates between the frames of the two stereoscopic view cameras, such that each of the pills/capsules displayed in a specified frame of one camera may be attributed with its counterpart at the second camera. The synchronized images are fed to a Decision Unit, DU, equipped with a Graphic Processing Unit, GPU, where the decision is made as to whether a major defect has been detected, or whether a critical event, such as the introduction of a foreign object, has occurred. Once a specific problem has

been detected, the DU sends the appropriate information to the Programmable Logic Controller, PLC, via Network Switch 1, which may cause the machine to stop or to reject the relevant bottle.

Classification Algorithm

The classification algorithm is based on a deep convolutional neural network (CNN) which is trained using images of good products, defected products and Mix-Ups. The training examples include recordings of machine run with good products, machine run with defected products, machine run with other types of products (Mix-Ups) and foreign objects (i.e. bolts and nuts) and several types of simulations.

“Deep Learning” (also called deep structured learning, hierarchical learning or deep machine learning) is a family of machine learning algorithms which is characterised by multiple processing layers. Since 2012, these methods have gained very high popularity in the field of computer vision and in other fields as well. Currently, the state of the art in many computer vision tasks is achieved with algorithms that are based on these methods.

The advantages of Deep Learning methods are:

1. They can generate very complex functions.

2. The classifiers can learn from examples without the need for manually engineered features.
3. They can be run and trained very efficiently on parallel computing platforms such as GPUs (graphical processing units).

The main disadvantages of Deep Learning methods are:

1. A large database of examples is usually required.
2. In most cases, the database should be manually annotated.
3. The system is a “black box”. It is hard to understand what the system has learned or to define specific criteria for the classifier. It is also hard to investigate failure cases.

These disadvantages make the use of deep learning for critical problems and major defects detection difficult. Specifically, the following are some issues that had to be overcome and some advantages accomplished by the invention:

1. In traditional methods, specific features are defined in order to specify the production tolerance i.e. to specify what is considered to be a defect. For example, the tolerance in tablet width or diameter, the tolerance in color, the maximal allowed fraction size, the maximal size of a spot on the product surface, etc. According to the invention no direct calculation of such features is employed. In

contrast, the invention relies only on examples (images) of good products, defected ones, Mix-Ups and foreign objects.

2. Capturing the products during free fall generates significant variability in their appearance. Specifically, in some cases overlapping products may appear. This would make it very hard to separate and measure features of each product (such as width, diameter, and shape) separately. The invention overcomes this hurdle by classifying the entire analysis region as a whole and by using many training examples including images of overlapping products. This approach has another advantage in run-time performance. Another cause of variation in appearance of the product while in free fall is the angle of the product, shadows and reflections that may appear.
3. Mix-Ups can result with a variety of foreign items, such as a different type of products, hard objects, such as nuts and bolts, etc. Accordingly, a critical problem may occur as a result of the existence of a Mix-Up that may be significantly different in appearance from the ones used in the training phase. The most challenging cases are when the classifier needs to detect a Mix-Up that was not learned during the training phase, and is very similar in appearance to the product being processed. According to the invention several types of simulations are used, including the generation of "hard examples" to cope with this issue. For example,

in a training phase a tablet that was very similar in appearance and only slightly different in size from the original processed tablet of the batch, was simulated by taking a segmented image of the original tablet and shrinking it by 10 percent, such that the overall appearance except the sizes was preserved. This simulation was then included in the data that was used for training and was marked as "critical problem".

4. There are tolerance differences in color and size between various batches of the same product. According to one embodiment of the invention the training scheme of the invention takes that into account by simulating inter-batch differences at training.

Generating a database with manual labelling of each product in the image is impractical. According to an embodiment of the invention a database generation scheme is used, which can generate millions of examples with none or almost no manual annotation effort. More specifically, in most real scenarios, the defected products or Mix-Up will appear close to other good products in the image. If a machine run with mixed products is recorded, a manual labelling of each product in the recorded images should be carried out. Instead, according to an embodiment of the invention separate recordings are used for each type of product, defect and foreign objects and simulations are used to implant mix-ups or defects (real or simulated) in a video of "good" products. Manual labeling is a tedious and time

consuming procedure. Rather, a simulated or segmented real image of a different product that was implanted deliberately to specified frames, is easy to label. This is illustrated in Fig. 8, which shows the image of a real tablet 81. The image of a capsule 82 was implanted in this frame by image processing, and this image was used to teach the system this possible mix-up. Similarly, the size of the articles can be changed by image processing (along with their color, which is not seen in the black-and-white figures) and the resulting image can be used for training. This is illustrated in Fig. 9 that has the real-life images 91 and 92 on the left side, and the corresponding changed images 91a and 92a on the right side. In one instance not only the size of tablet 91 was changed, but also its color was changed from blue (tablet 91) to pink (tablet 91a). However, this additional change is not seen in the black-and-white figures.

5. Though stereo images are used in some embodiments of the invention, there is no need to compute explicit 3D information or 3D reconstruction. According to the invention the classifier is fed with images of both cameras so it can learn to use the 3D information for classification without an explicit 3D reconstruction. This approach can be generalized for configurations with more than two cameras. It is advantageous to use two cameras in order to have a larger coverage of the underline inspected pill/tablet, that significantly increases the detection of defects and of critical events.

Training procedure

The training procedure according to an embodiment of the invention comprises the following steps:

1. Record a machine run of good products.
2. Record a machine run with defected products.
3. Record machine runs with Mix-Ups and/or use previously recorded images of other products and foreign objects.
4. Activate Foreground-Background segmentation to detect the pixels which contain products. The segmentation generates blobs that contain the product's pixels.
5. Choose frames for training: a. all of the images of the good product record. b. from the Mix-Up and defected product records, choose frames where a significant portion of at least a single product overlaps the analysis region of interest (ROI). This is done by analysing the area and location of the blobs computed by the foreground-background segmentation. This step is done to prevent the use of ambiguous frames in the training phase. An ambiguous frame is a frame which does not contain enough information (a large enough portion of the product) to enable unique classification. The training database can be further filtered to detect ambiguous or misclassified images in the training database. Further details are provided in the "Filtering the Training Database" section below.

6. Augment the database using simulations:

- a. Introduce random shifts to the images to simulate minor shifts of the cameras. This increases classifier robustness to camera shakes\movements. For example, a slightly tilted camera relative to its original position introduces a shift of the tablets/capsules in the field of view, relative small shift to the background and slightly different illumination effects. All these changes may be simulated using the original captured frames as a basis data for the change and then be added to the training data.
- b. Introduce minor changes to the light color to simulate minor changes to the lightening conditions. This increases robustness to lighting changes. A second mechanism to compensate for light changes is the use of a calibration target that is visible in the field of view and is used to compensate for light changes.
- c. Introduce a random color change, size change and distortion change to the Mix-Up products blobs to simulate more types of Mix-Ups.
- d. Introduce a random color change, size change and distortion change to the good products blobs to simulate other types of Mix-Ups and to generate "Hard Mix-Up Examples" for training. i.e. generate images of products with only small

change of size or color (the change must be larger from the intra batch tolerance).

- e. Implant Mix-Ups and defected blobs (either real or simulated) in images of the good record to generate additional training images for the relevant classes.
- f. Introduce minor color change and possibly minor size change to good product blobs. This increases robustness to inter-batch differences.
- g. Background subtraction \ darkening – it is possible to darken or to replace or change the background in order to be robust to changes in background that may be introduced by machines that are different from the one used for training.
- h. It is possible to use additional simulations of Mix-Ups by using rendering of 3D models of products or foreign objects.

The database generation scheme according to one embodiment of the invention obviates the need for manual annotation of the DB (Data Base).

In other words, if a machine run is recorded with mixed products, then a human must later annotate the video for training the classifier, i.e., the human must annotate each frame of the video as good\major\critical. The invention obviates the need for this stage

by recording a machine run with a single type of product and simulating images with several types of products.

The simulations generate any desired example of mix-up and defects, including hard-to-detect examples (i.e., cases in which it is hard to distinguish between a correct product and a Mixed-Up product because they only slightly differ in color, shape and size). It further implants defects and Mix-Ups near good products, which match real-life scenarios. Recording these "real cases" is not practical for a large database since such videos should be manually annotated later, resulting in a staggering amount of human manual work.

7. Train a deep convolutional neural network to classify each frame (or each series of consequent frames) in the training set as either Good \ Major \ Critical. In one embodiment, the network was fed with a stereo pair, i.e. for each channel of the tray, the ROIs from the two cameras were taken and aligned side by side, forming a single (stereo pair) image for that channel. In this exemplary embodiment the network architecture used was that described in, Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Going deeper with

convolutions." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-9. 2015, although other architectures may be used. The classification layers were replaced (this architecture contains 3 classification layers in the training phase. Only the last classification layer is used in the testing phase) with new classification layers which contain 3 neurons matching the 3 classes (Good\Major\Critical). The network coefficients (excluding the coefficients of the classification layers) were initialized with a network that was trained on ImageNet classification challenge (as described in Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. *IJCV*, 2015). The classification layers coefficients were initialized with a random Gaussian distribution. The back-propagation algorithm was used to train the network, as described in "Pattern Classification" by Duda, Hart, Stork, Wiley & Sons Inc., 2001, pp.288.

Filtering the Training Database

1. The defected area of a defected product may not be visible in all of the frames of the "major defects recordings" which are used in the training phase. This may cause problems in the training. For

example, a break might not be visible in a specific stereo pair. In such case, on the training phase, the training algorithm is fed with the stereo pair and with the “major defect” label while no defect is visible in the frame. Such training examples may adversely affect the accuracy of the resulting classifier. This problem is mitigated if the classifier uses several frames for classification as there is a much lower chance that the defect will not be visible in all of the sequence frames.

2. Another problem may be caused by mistakes in the training DB. For example, a defected product may exist in the “Good” recording or a good product may exist in the “major defect” recording.
3. To cope with these issues, the training procedure may contain a filtering step in which ambiguous frames or misclassified frames in the training set are identified and are either ignored or given a corrected label. This filtering step may be done automatically or semi-automatically. One strategy is to divide the Good and Major products into two groups. The recording of each group is used to train a CNN. Then, the label of each image of the first group is evaluated by the CNN which was trained on the second group and vice versa. Frames which are misclassified or with score (for the correct class) that is lower than a threshold are presented to a human observer for evaluation of their true class (the human observer may also decide that the frame label is ambiguous).

4. Since in most cases, the defect is visible in at least one of the frames of the product fall, the classifier for the DB filtering stage may use this assumption to improve its robustness. For example, a multi-frame classifier for the filtering stage may use a "Siamese CNN" architecture in which the network input is a sequence of frames (say 4). Each frame in the input sequence is fed to the same network (therefore termed "Siamese"). The network generates a "Major" score and a "Good" score for each of the input frames (both scores are between 0 and 1 and their sum is 1). The loss function who guides the optimization of the CNN during its training may rely on the maximal "Major Score" of frames in the sequence. This scheme has the advantage that if the defect is not visible in some of the frames of the input sequence, it will not affect the classifier training as long as the defect is visible in at least one frame of the sequence. After training of this "Siamese CNN", it may be used to calculate a score for a single frame so the DB filtering procedure defined in the previous paragraph may rely on this classifier.

Another strategy to reduce the sensitivity of the main classifier (not the DB filtering classifier) to ambiguous frames in the major recordings is to use a "multi frame" classifier. In most cases, the defect is visible in at least one of the frames of the product fall. Therefore, if the network uses several consecutive frames for classification it can reduce the sensitivity of the

training to the filtering of the major recordings. It might even make the DB filtering step unnecessary in some cases. This strategy may be combined with a database filtering step as described in the previous paragraphs.

In inspecting time

The following steps are performed during inspection:

1. Crop the analysis region of interest from the image.
2. Darkening or replacing the background – optional.
3. Feed the image into the neural network and get a score for each of the classes: Good \ Major \ Critical.
4. In some embodiments, the classifier may use several frames for classification.
5. Custom decision logic may be defined to decide on action depending on the scores of the classifier. For example, one may decide that if the good score is smaller than some threshold T then the machine will either stop or reject a bottle, depending on the score of the critical error. A different logic is WTA (winner takes all) meaning that the highest score determines the action i.e. if Critical is highest then the machine will stop. If Major is the highest, then the machine will reject a bottle. Sensitivity parameters may also be used to multiply the scores of the classifier with some pre-defined parameters or to do some other mathematical manipulation of the scores. The system may be

configured to save some of the frames for future investigation. For example, if a critical or major event occurred or if the good score is smaller than some pre-defined threshold. (Other cases when the images should be saved may be manually defined.)

6. Count the products which passed through the analysis region by using foreground-background segmentation. The segmentation algorithm uses methods commonly known in the art for foreground-background segmentation, such as the use of background modelling and background subtraction.

According to an embodiment of the invention, the items in the image are not individually identified and, instead, the image is analysed in its entirety. An image may contain any number of items, which may also overlap and hide one another, partially or fully.

The decision logic and sensitivity

The sensitivity and decision logic may be defined manually. Alternatively, these parameters may be defined automatically by defining the cost of each type of mistake. There are 6 types of mistakes:

Classifying a good product as major;

Classifying a good product as critical;

Classifying a major product as good;

Classifying a major product as critical;

Classifying a critical product as good;

Classifying a critical product as major.

If the cost of each type mistake is defined and some estimation of the frequency of each class (priors) is available, then a Bayesian decision approach [see, e.g., James O. Berger, "Statistical decision Theory and Bayesian Analysis", 2nd Edition, Springer-Verlag: <https://www.math.ntnu.no/~ushakov/emner/ST2201/v08/files/berger1.pdf>] may be used as decision logic in order to minimize the estimated cost of the mistakes. The decision parameters may be calibrated using videos generated by simulations. According to one embodiment of the invention this method is used to decide which action to take, among the possible actions: 1. continue, 2. stop the machine 3. reject a bottle, in light of the computed classification scores. However, other decision methods can also be employed.

All the above description of preferred embodiments have been provided for the purpose of illustration and are not intended to limit the invention in any way. Many variations, types of neural networks, training schemes and different apparatus can be used, without exceeding the scope of the invention as defined in the appended claims.

Claims

1. A method for inspecting and classifying discrete items, comprising capturing the images of a free-falling item using at least one imaging device and analyzing the captured images using a neural network which has been trained using images and/or simulated images of regular products, defected products and Mix-Ups.
2. A method according to claim 1, wherein the item's free fall occurs from a filling machine.
3. A method according to claim 1, wherein images from two or more imaging devices are fed to the classifier.
4. A method according to claim 1, wherein the discrete item is a pharmaceutical or nutraceutical product.
5. A method according to claim 1, wherein the classification includes items that are "Good", have a Major Defect, or have a Critical Problem.
6. A method according to claim 2, wherein if a Critical Problem is detected the operation of the filling machine is stopped.
7. A method according to claim 5, wherein if a Major defect is detected, one or more packaging that have been filled during such detection are removed from the production line.
8. A system for inspecting and classifying discrete items, comprising at least one imaging device suitable to capture the images of a free-falling item, and analysis apparatus suitable to analyze the images

captured by said imaging device to assign a class to it according to a preset classification.

9. A system according to claim 8, wherein one or more image device(s) are suitable to capture the free-falling article before its free fall begins.
10. A system according to claim 8, wherein the analysis apparatus comprises a Decision Unit provided with a Processing Unit and wherein said Decision Unit operates a neural network.
11. A system according to claim 8, which is provided with a database including data generated by simulating a change in properties of items in the captured images.
12. A system according to claim 11, wherein the change in properties comprises a variation in size and/or color and/or shape and/or background of an item, generated from a real picture thereof.
13. A system according to claim 10, wherein the items are selected from tablets, pills and capsules, including chewy medicines.
14. A system according to claim 8, comprising rotation-imparting elements suitable to cause the item to rotate during its fall thereby permitting an extended view and inspection of the surface thereof.

15.A method for training a classifier for classifying images or sequences of images of items comprising the steps of:

a) capturing a sequence of images which contain only a single type of either good or defected items or Mix-Ups;

b) segmenting said images using foreground-background segmentation;

c) generating additional images using one or more of the following options, alone or combined:

1) implanting segmented blobs of items from images of a specific type of items (either defected or Mix-Ups including simulated Mix-Ups) into images of "Good" items, in order to create images which contain multiple types of items;

2) simulating changes of lighting of the images;

3) simulating changes in the background of the items, including replacing the background with a constant background or replacing the background to match a background of a different machine or randomly change to the background;

4) simulating several levels of changes to the Good items shape, size and/or color including minor changes which are in the allowed tolerance of the "Good" items, and larger changes which exceed said allowed tolerance, including changes that slightly exceed said tolerance;

5) Simulating changes to Mix-Ups size and/or color and/or shape.

5) simulating minor shifts of the cameras;

6) generating images of rendered 3D models of items;

7) implanting rendered 3D models of items in captured images;

and

d) using said captured and generated images to train a classifier for classifying images or sequences of images.

16.A method according to claim 15 further including the step of filtering the training database by removing from the training DB images of the non-good classes recordings in which a defect is not visible or no article is present in the image and also removing images from the good record which might contain a defected product by one or more of the following:

1) using a foreground-background segmentation to analyze the presence, area and location of items in the images and removing from the training DB images of the non-good records in which no product is presence in the image or only a small portion of the product overlaps the analysis ROI.

2) training a classifier to distinguish between images from the good record and images from the defected records and removing

from the training DB images which got low score for the class which matches the record they were taken from.

3) same as in 2), but the DB is split into two sets and a separate classifier is trained on each of these sets and is activated on the second set in order to get a classification score for each of the classes.

4) same as in 2) or in 3), but the classifier is trained using "Siamese" architecture and is fed during the training with sequences of frames in order to improve robustness.

5) same as in 2) or in 3) or in 4), but a human observer is asked to confirm the removal of an image from the training DB.

1/9

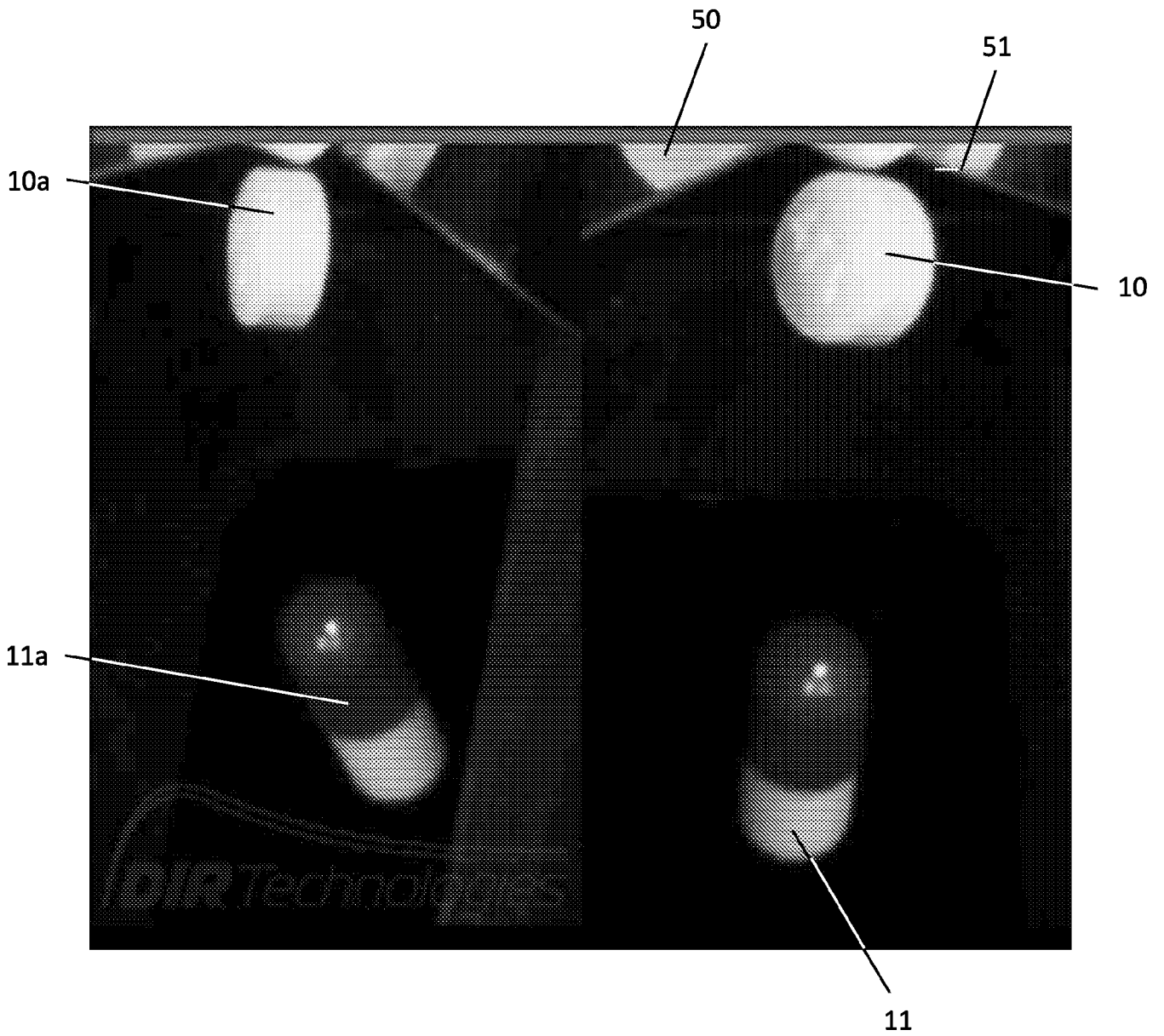


Fig. 1

2/9

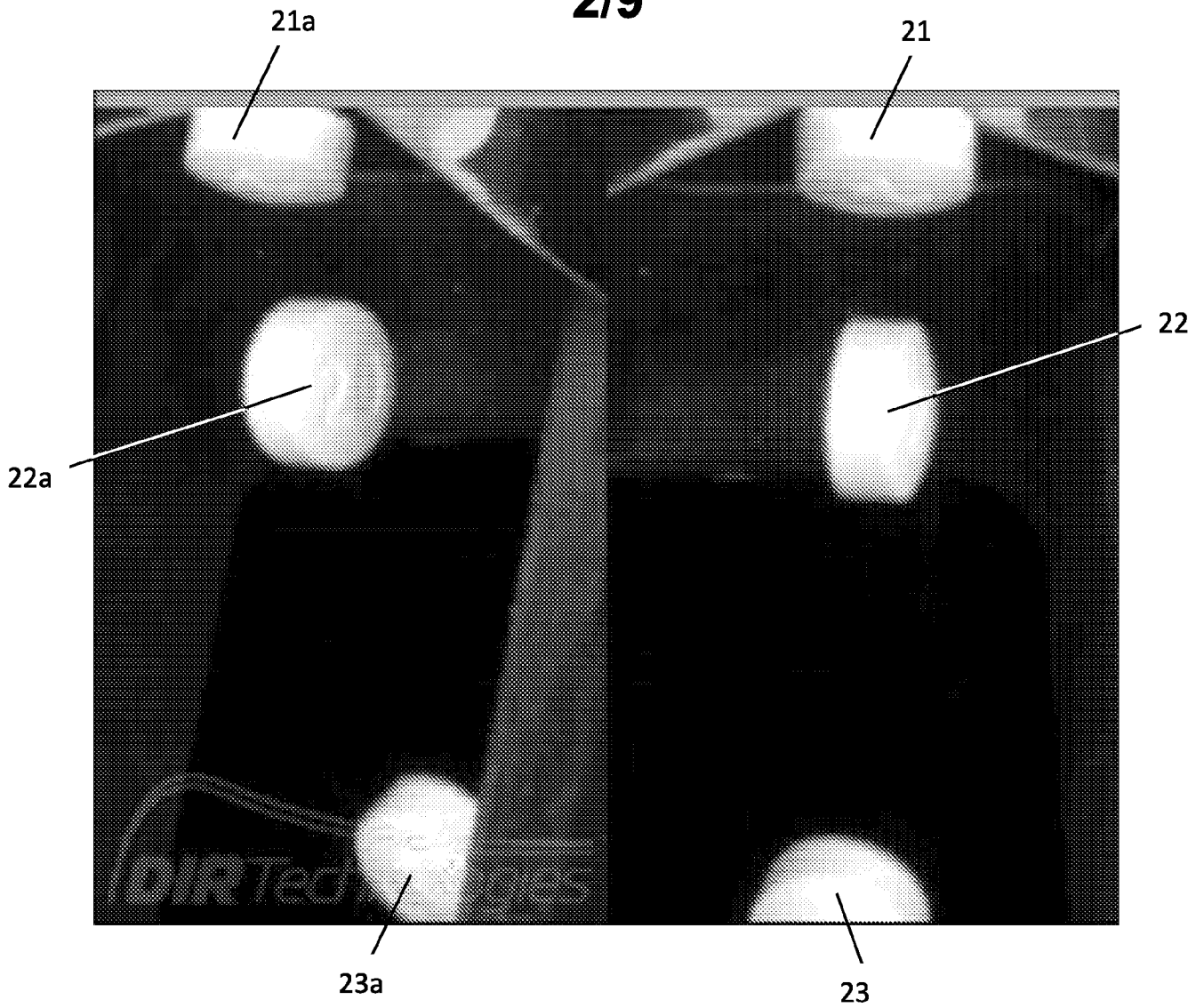


Fig. 2

3/9

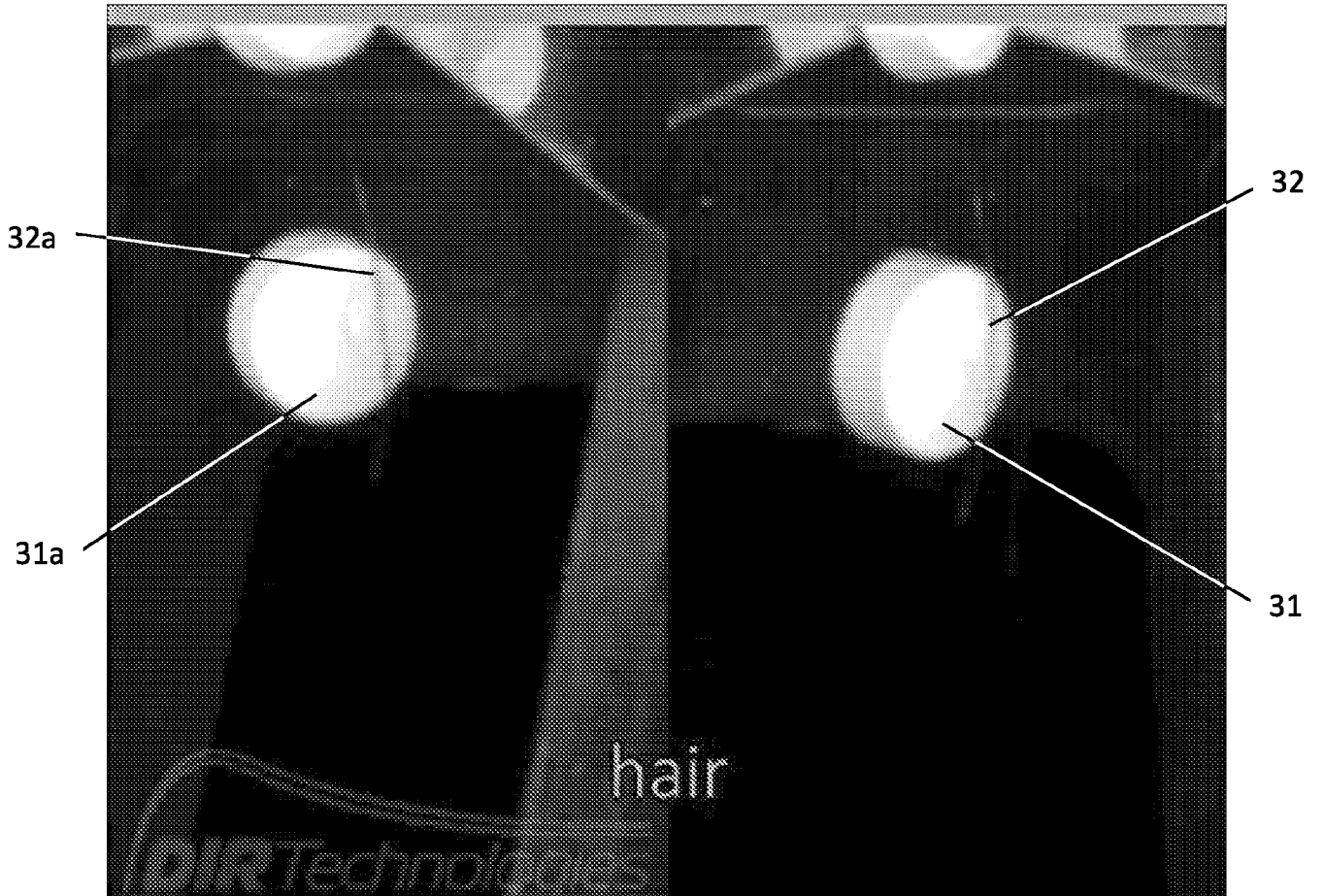
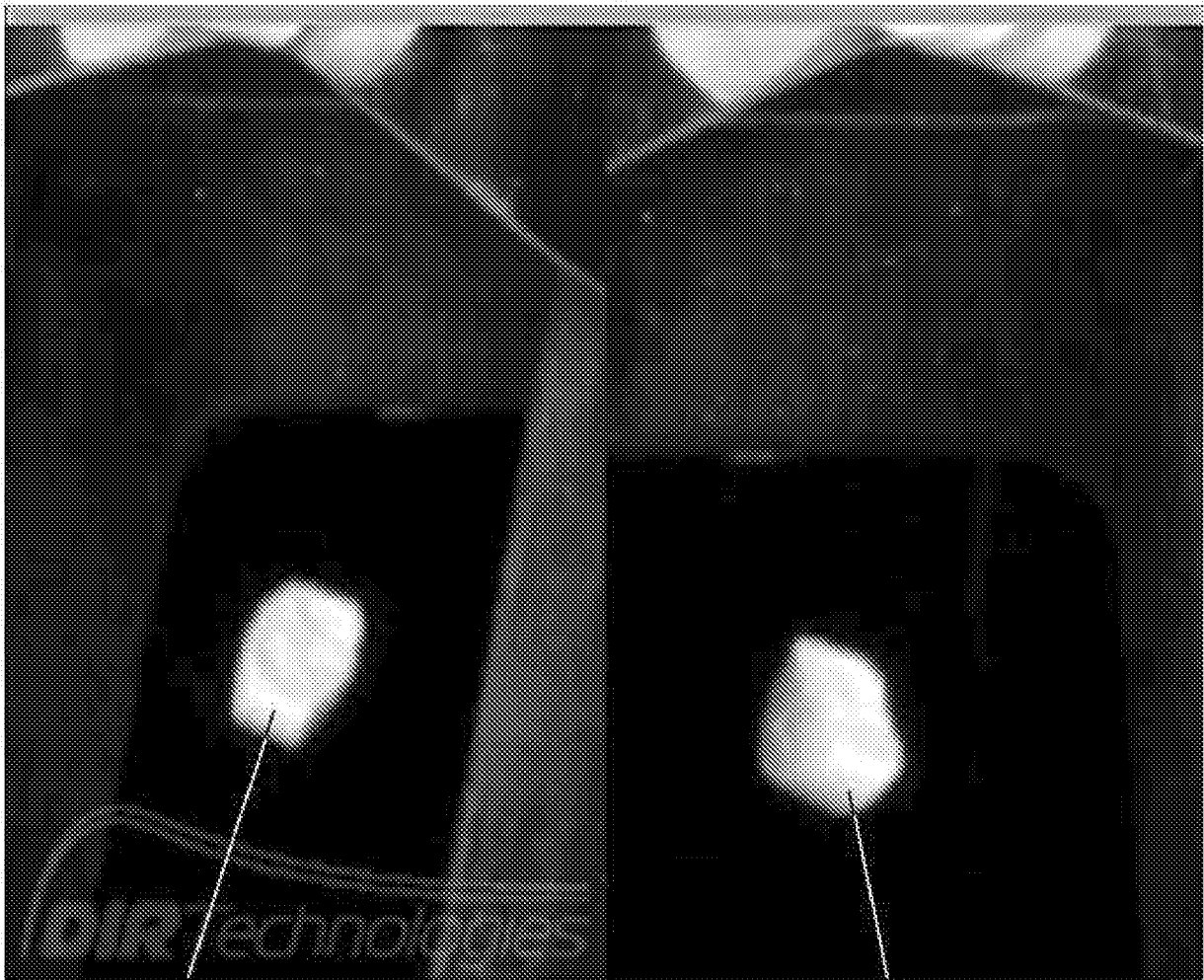


Fig. 3

4/9



41a

41

Fig. 4

5/9

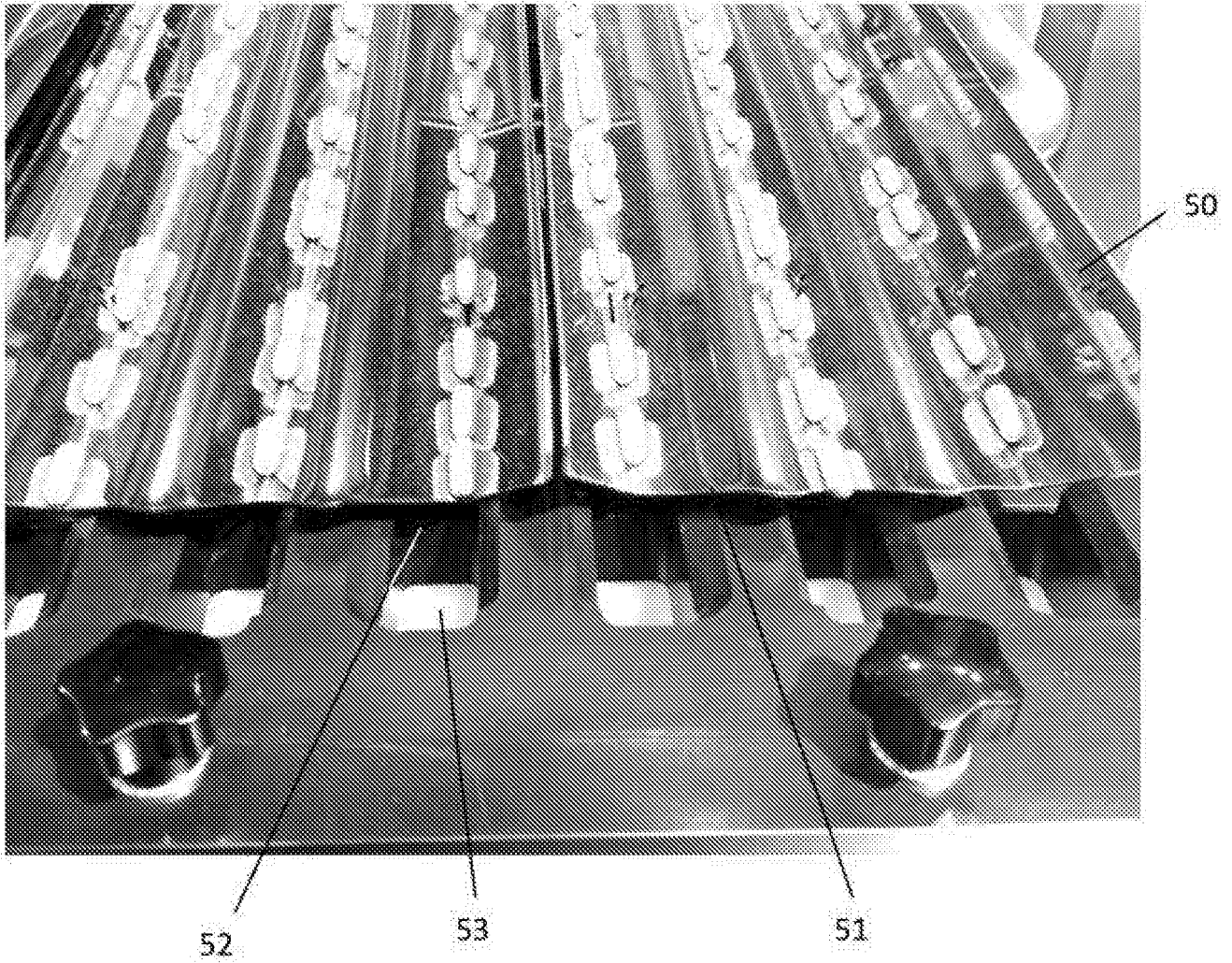


Fig. 5

6/9

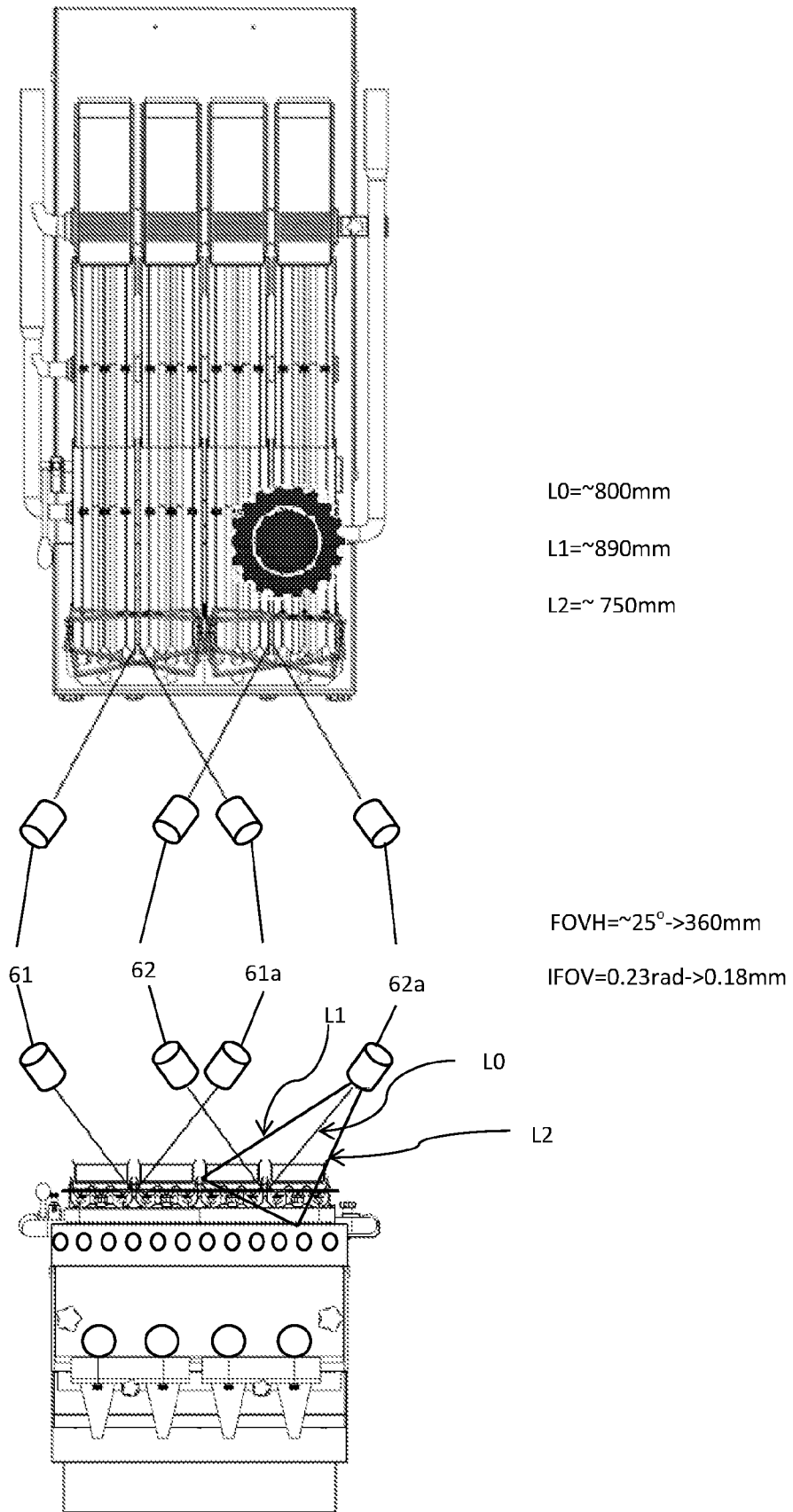


Fig. 6

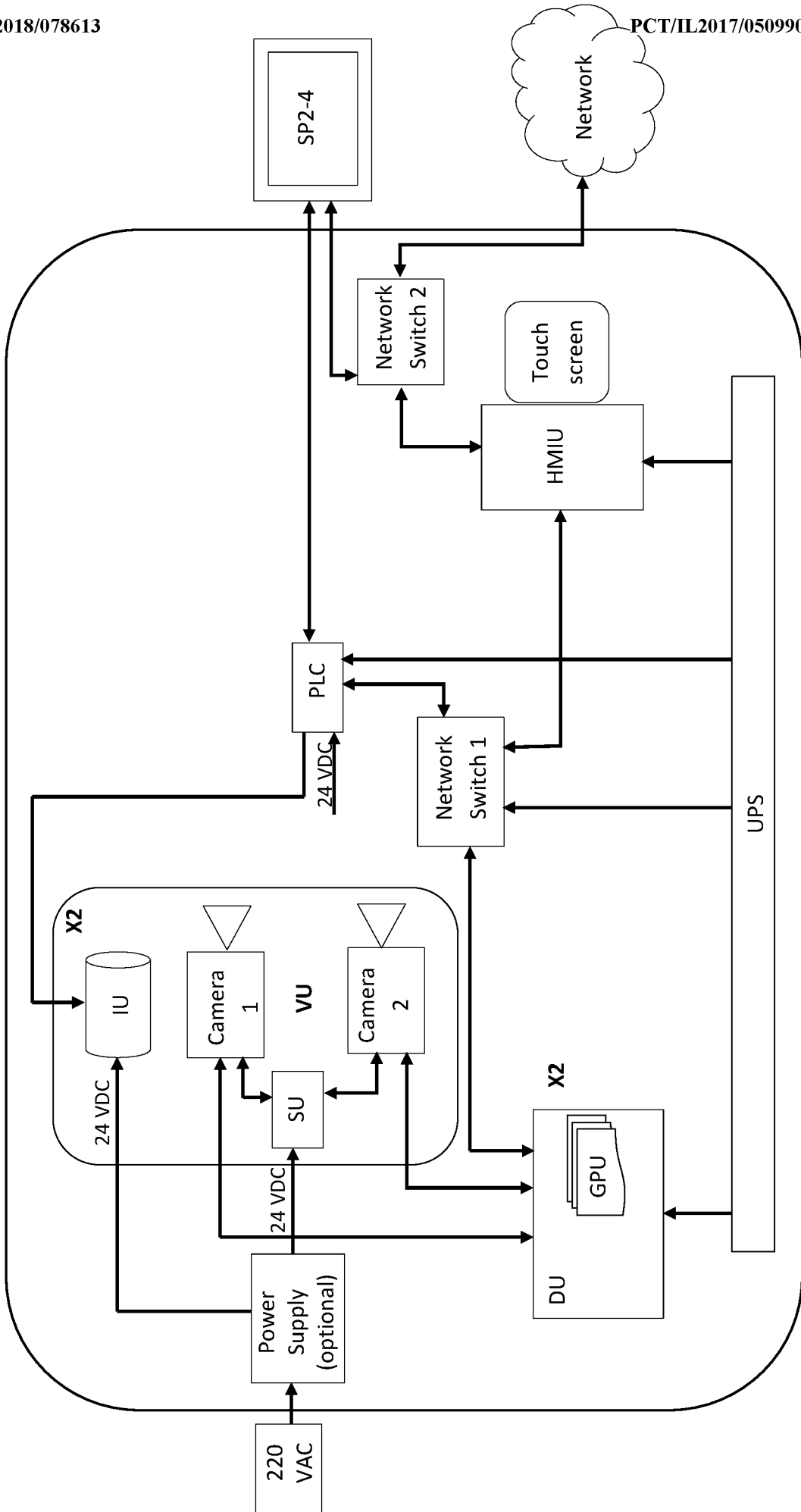


Fig. 7

8/9

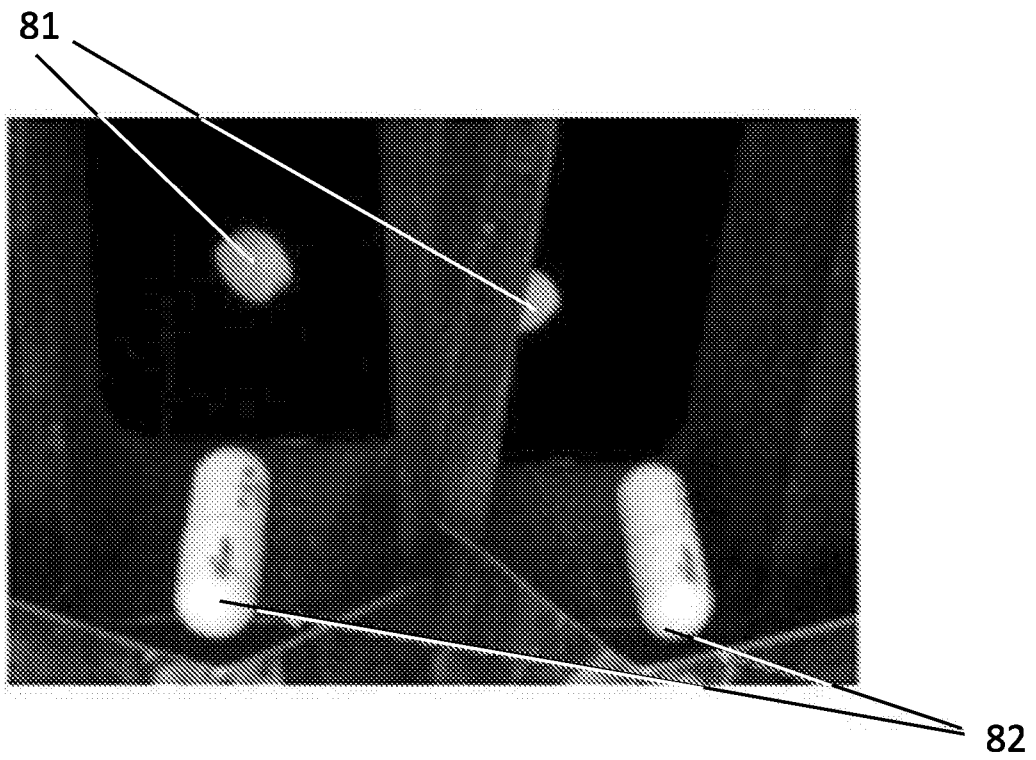


Fig. 8

9/9

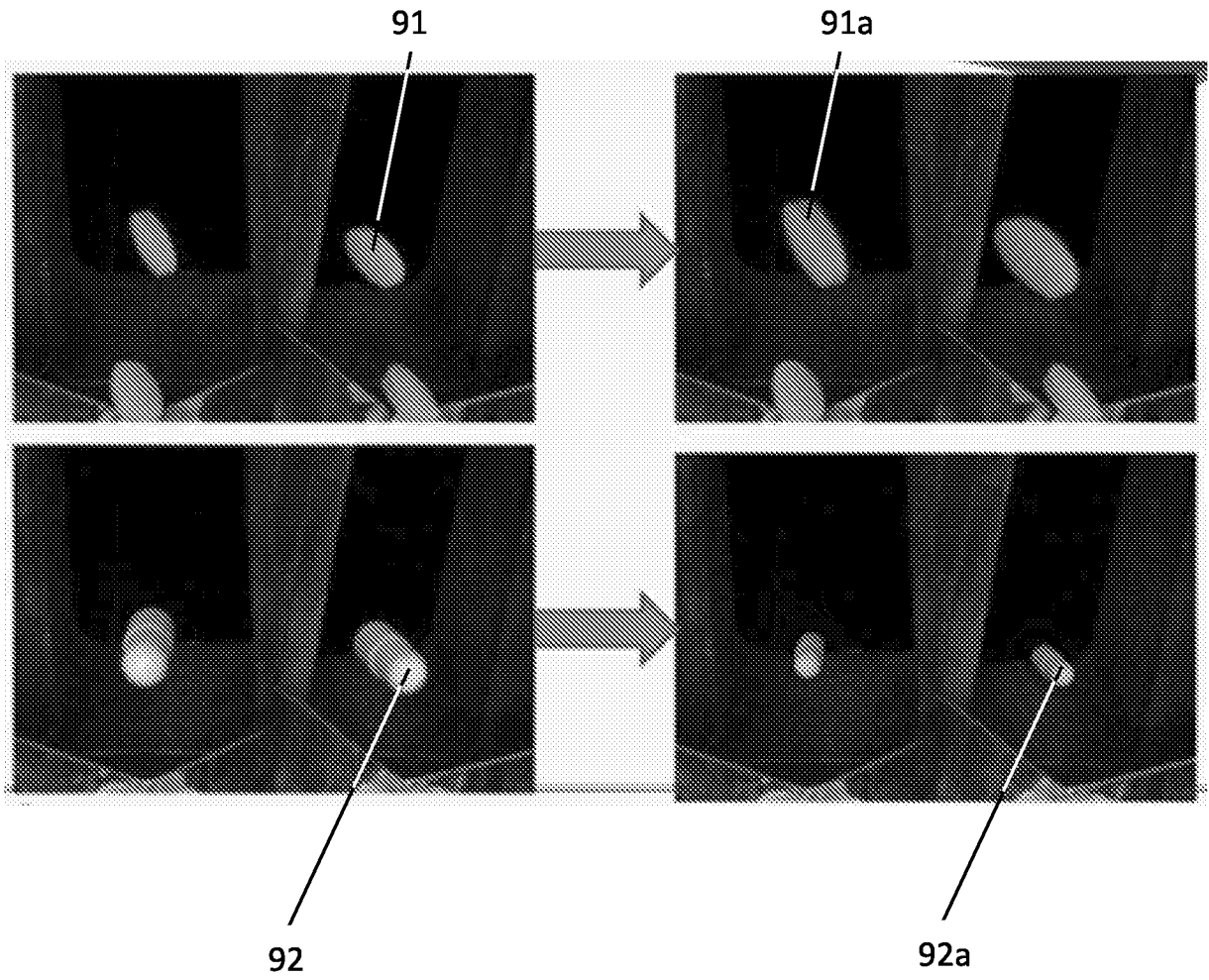


Fig. 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IL2017/050990

A. CLASSIFICATION OF SUBJECT MATTER

IPC (2017.01) B65B 57/14, G06K 9/46

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC (2017.01) B65B 57/14, G06K 9/46

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Databases consulted: Esp@cenet, Google Patents, Google Scholar

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Article: Machine Vision based Quality Control: Importance in Pharmaceutical Industry. IJCA Proceedings on International Conference on Information and Communication Technologies ICICT(8):30-35, October 2014 [retrieved on 2017/11/01]. Retrieved from http://research.ijcaonline.org/icict/number8/icict1491.pdf Gundarapu Pavan Kumar and Pinnamaneni Bhanu Prasad 01 Oct 2014 (2014/10/01) whole document;	1-16
Y	Research on Defect Detection Technology of Tablets in Aluminum Plastic Package. The Open Automation and Control Systems Journal (6): 940-951, 2014 [retrieved on 2017/11/02]. Retrieved from https://pdfs.semanticscholar.org/1f25/ef84c7bdd88ac22251e4a2faae98d2ca98eb.pdf . Yueqiu Jiang et al 01 Nov 2014 (2014/11/01) §5, 6	1-16

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

03 Dec 2017

Date of mailing of the international search report

05 Dec 2017

Name and mailing address of the ISA:

Israel Patent Office
Technology Park, Bldg.5, Malcha, Jerusalem, 9695101, Israel
Facsimile No. 972-2-5651616

Authorized officer

MARCOWITZ Noam

Telephone No. 972-2-5651735