Automatic Inventory of Multi-part Kits Using Computer Vision

A. J. Rodríguez-Garrido, A. Quesada-Arencibia(✉),
J. C. Rodríguez-Rodríguez, C. R. García, and R. Moreno-Díaz jr

Institute for Cybernetic Science and Technology,
University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain
jgadir2000@gmail.com, {alexis.quesada, ruben.garcia,
roberto.morenodiaz}@ulpgc.es, jcarlos@iuctc.ulpgc.es

Abstract. A prototype tool for the detection, segmentation, classification and counting of Lego pieces based on the OpenCV artificial vision library is presented. This prototype arises before the need to automate the complex and tedious task of the inventoried one of Lego kits of the MindStorm serie.

In the process of detection and segmentation there have been used skills of threshold and the algorithm of Watershed segmentation. For the process of classification and count have been used two different approaches in the securing of the vector of characteristics of the image: BOW and Naive; as well as vector machines support (SVM) for the classification.

Keywords: Automation · OpenCV · Lego · Prosecution of images
Segmentation · Extraction of characteristics · Classification

1 Introduction

The intention of our work is to automate the classification and re-count of Lego pieces. The MindStorms serie of Lego is a powerful tool for the educational robotic [1], the conservation of the material in good conditions involves a tough logistics, which completes time and energy, and which nevertheless is necessary. The action to count pieces allows to determine when there are or not missed pieces or if we have all the pieces necessary for the class.

The solution that is proposed belongs to the field of the artificial vision. One works with images of scenes that contain interest objects, in our case Lego pieces of the kit. We must be capable of detecting them, of classifying them and of counting them. This process of visual perception must acquire images, detect and segment the interest objects, and finally classify them and count them (to see Fig. 1).

The solution that is proposed is characterized by a modular structure, represented by a module of "Acquisition of Images" of physical nature and two modules of nature clearly software that there shape a library or bookstore (modules of "Detection and segmentation" and "Classification and count").

R. Moreno-Díaz et al. (Eds.): EUROCAST 2017, Part I, LNCS 10671, pp. 142–149, 2018.
https://doi.org/10.1007/978-3-319-74718-7_17

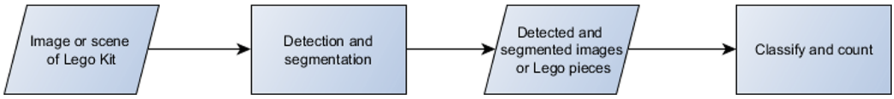

Fig. 1. Process of visual perception.

2 Images Acquisition: The Module of Acquisition of Images

For our intention, it is necessary to be provided with some device that should allow to realize image[1] apprehensions with the Lego pieces placed in a qualified space.

In this respect, the optical device will have to assemble a few certain characteristics in order to guarantee the quality of the above mentioned apprehensions. Therefore, at the time of praising us for an option or other one we will have to value aspects as the space of color that we are going to need (gray or of color), resolution, type of sensor (CCD or CMOS), focal distance (a lens adapted to its place and Surface of work), … For this intention, we have integrated the following elements:

- Kaiser Platform in which the Lego pieces are placed of the kit to audit. Ideal work condition: Surface of 35 cm^2 and 50 cm focal distance.
- Webcam[2] of Logitech connected to a terminal type PC where the apprehensions are stored and with facility for the remote access and distribution of the same ones.

Fig. 2. Area of work delimited by a rectangle.

In order to assure the maximum level of detail of the image, we establish the following working conditions (to see Fig. 2):

- Images in scale of gray (major contrast).
- 43.9 cm focal distance.
- Surface of work of 8.5 cm^2

[1] Examples of optical devices are a camera, video camera, Ethernet camera, web-cam…

[2] Initially it was decided in favor by the Mako high resolution Ethernet camera. However, on not having counted with optical adapted to be able to work with the mentioned camera, it was necessary to use an optical device of worse resolution, the Logitech webcam.

3 Detection and Segmentation

3.1 The Reprocess

Once we have obtained for some of the devices described in the previous paragraph the set of images to be processed, the following step is to be able to detect and segment to (extract) each of the objects from interest (Lego pieces of the kit) presents in the above mentioned images.

For it, it is necessary to realize previously a preprocess of the image, dependent on the context in which the apprehensions have been realized and that allows us to adapt it to the different algorithms or skills of segmentation.

This preprocess is usually intimately tied to the environment or context in which the acquisition of the images is realized. In our proposal, it is a specific procedure, of manual and key adjustment, since it intervenes both in the securing of the scoreboards that uses our segmentation process is realized in a controlled environment, there are used the following skills of preprocess of image (to see Fig. 3).

- Conversion to scale of gray (space of color needed by the elected algorithm of segmentation).
- Equalization of the histogram of the image (it improves the contrast of the image).
- Morphologic Operation of dilation (it eliminates hollows or spaces inside the objects).
- Inverse binary threshold, using like value of threshold obtained by the method of Otsu [2].

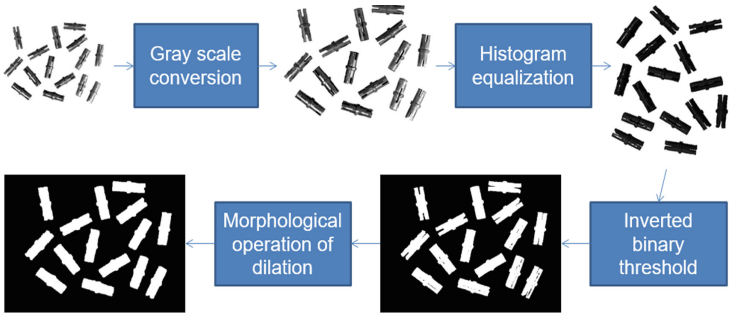

Fig. 3. Steps of the preprocess.

3.2 The Segmentation

The segmentation allows us the extraction and therefore detection of one or more regions or objects of interest of the image (Lego pieces) base on a criterion of discontinuity or similarity. This is; to go on from an image or scene containing the Lego pieces of the kit to be audited to a set of Lego objects or pieces detected and extracted (segmented) in that image. This process in divided into:

1. Securing of the scoreboards of our algorithm of segmentation.
2. Application of our algorithm of segmentation to the image or scene.

As segmentation algorithm, we have used the algorithm of segmentation based on regions Watershed [3, 4] for its confirmed hardiness and the results obtained in our preliminary tests. This algorithm considers an image in scale of gray like a topographic relief, where the level of gray of a pixel is interpreted as its altitude in the relief.

Although the segmentation process concludes with the securing of the set of detected and abstracted images, there has been included the option to normalize the above mentioned set of images, that is to say, all resulting images are of the same size. This decision justifies to the being an aspect that influences in the behaviour and yield of the process carried out by the Classification module and count.

The flow of the process is described in the Fig. 4 and can be appreciated in the Fig. 5.

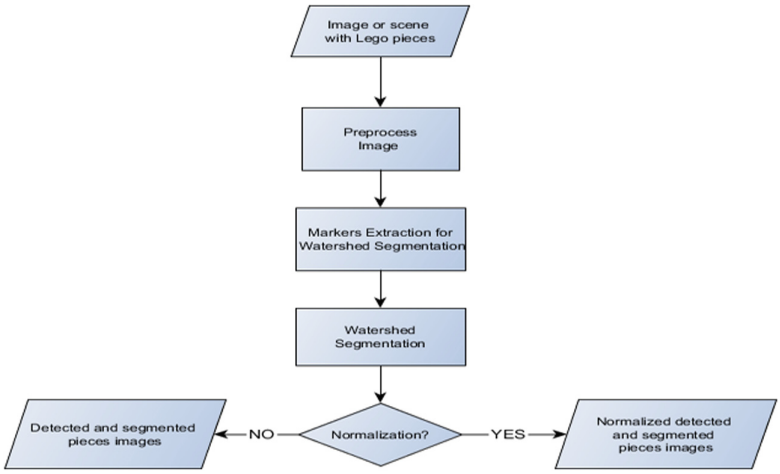

Fig. 4. Flow of the process carried out by the detection and segmentation module.

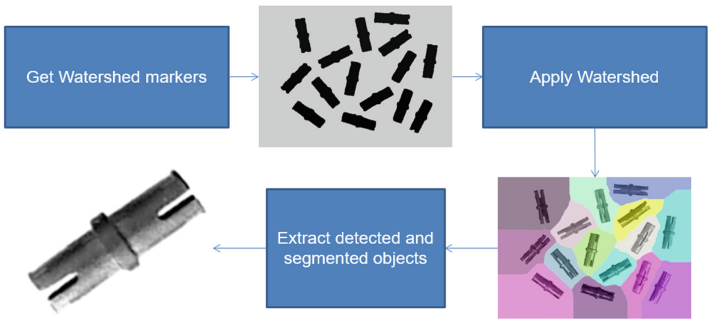

Fig. 5. Different stages of the segmentation.

4 Classification and Count

We come to the module entrusted to realise the process of classification and count of the Lego pieces of the kit to be audited. We will spend of the set of normalized images obtained in the stage before to a list containing the number of pieces for category of the kit recognized, strangers and those that were not possible of classifying to more than one category to be assigned.

This process consists of the phase of securing of vectors of characteristics from the descriptors of image and the phase of classification and count in the strict sense.

4.1 Phase of Securing of the Vector of Characteristics

Image descriptors

The images descriptors are mathematical gadgetry that allow us to identify of univocal form an image, in spite of presenting transformations due to changes of lighting, noise, scale, rotation… Bearing in mind the need for visual invariability opposite to rotations (pieces interspersed in arbitrary position) our solution he supports the use of the following descriptors of images:

- Dense detector and like extractor well SIFT or SURFING
- KAZE and AKAZE
- Moments of HU [5]

This support flexibility allows to be studied by the descriptor who better adapts himself to our problem.

SIFT [6] and SURFING [7] are examples of descriptors based on histogram of gradients of image (direction and magnitude of every pixel). There is an invariant opposite to changes of lighting, but they do not capture geometric information.

KAZE [8] and AKAZE [9] are descriptors who have appeared recently. They are characterized for using not linear spaces of scale which, unlike previous descriptors, respect the natural outline of the objects.

Approaches

For this phase two approaches were used (to see Fig. 6):

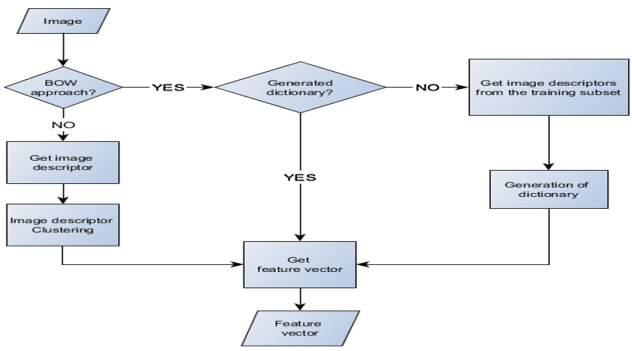

Fig. 6. Flow of the process carried out by the phase of securing of characteristics.

Confirmed standard approach
There is used the generative model BOW [10] to generate the characteristics vector. On having used a words dictionary, we obtain a major standardization of the results and reduce the redundancy in the training information. The construction of the dictionary is realized from the obtained descriptors of images of a subcommittee of images of training.
"Naive" experimental approach
The characteristics vector is obtained from a clustered of the descriptor of image. The main motives of using this approach are both its simplicity and theoretically a less training time with regard to the previous approach. This approach is one of the original contributions of this work.

4.2 Phase of Classification and Count

To the being the mastery of our problem well-known, finite and limited size, we use like classifier the algorithm of supervised learning SVM [11]. Therefore, a phase previous to training is needed. The training information will be obtained from the set of vectors of characteristics associated with the set of images of normalized[3] training of every SVM or classifier.

As soon as the set of sorters was trained (as many as categories of the Lego kit), we proceed to the classification and count of the Lego pieces of the kit to audit. For it, from the vector of characteristics associated with every piece, we obtain the result of the-classifiers. Based on the obtained result we will proceed to increase in a unit (to see Figs. 7 and 8) the value of:

- the category that represents it if only a classifier identifies it.
- the special category "doubtful" if more than one classifier identifies it.
- the special category "known" in case of not being identified.

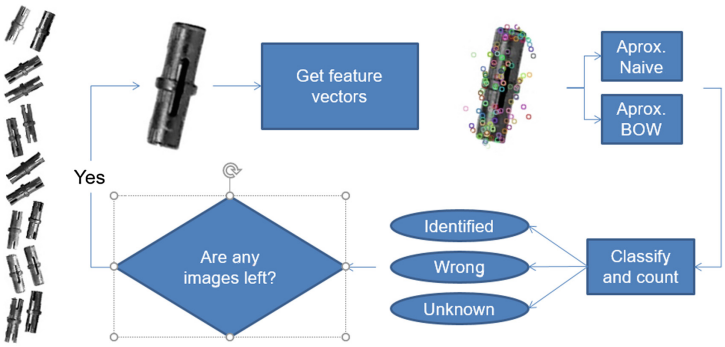

Fig. 7. Diagram of the classification and count.

[3] It is important that both sets of images (training and classification) are normalized, that is to say, the images that integrate it are of the same size. In the opposite case, we might obtain vectors of characteristics that they do not represent from trustworthy form to the above mentioned sets.

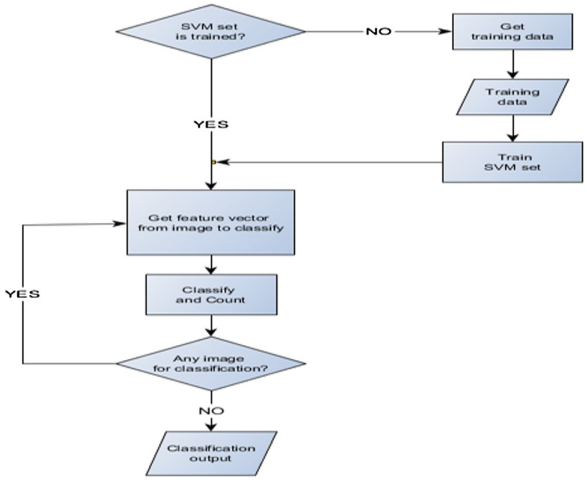

Fig. 8. Flow of the process carried out by the classification phase and count.

5 Contributions

A vision system has appeared of artificial highly configurable, with two approach ways for the classification and count, one of which has been proposed by the authors.

Between others they are used of binary images descriptors novel like KAZE and AKAZE not supported completely at the moment of the design and implementation of the described solution.

As a result of the previous thing we have:

- The possibility of using descriptors of images not supported by the bookstore OpenCV [12] whenever its interface is respected.
- The implementation of a skill of binary clusterization k-medoids, as well as its later use on the part of two approaches of classification and count.

This approach proposed by the authors or "Naive" approach, it is provided with a high valuation of reliability (number of pieces correctly identified) according to the results of a wide battery of tests of combination of methods (procedure used for the securing of the vector of characteristics, image descriptor) and configuration (size of the cluster, type of nucleus SVM). The results, as well as its conclusions, the authors hope to see it published in a future article.

References

1. Rodríguez, J.C.R., Martín-Pulido, E., Padrón, V.J., Alemán, J.A., García, C.R., Quesada-Arencibia, A.: Ciberlandia: an educational robotics program to promote STEM careers in primary and secondary schools. In: Auer, M.E., Guralnick, D., Uhomoibhi, J. (eds.) ICL 2016. AISC, vol. 544, pp. 440–454. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-50337-0_42

2. Otsu, N.: A threshold selection method from gray-scale histogram. IEEE Trans. Syst. Man Cybern. **9**, 62–66 (1979)
3. Beucher, S., Lantuejoul, C.: Use of watersheds in contour detection. In: International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation (1979)
4. Meyer, F.: Color image segmentation. In: International Conference on Image Processing and its Applications (1992)
5. Hu, M.K.: Visual pattern recognition by moment invariants. IRE Trans. Inf. Theory **8**, 179–187 (1962)
6. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**, 91–110 (2005)
7. Bay, H., Ess, A., Tuytelaar, T., Van Gool, L.: Speeded-up robust features (SURF). Comput. Vis. Image Underst. **110**(3), 346–359 (2004)
8. Alcantarilla, P.F., Bartoli, A., Davison, A.J.: KAZE Features. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012. LNCS, vol. 7577, pp. 214–227. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33783-3_16
9. Alcantarilla, P.F., Nuevo, J., Bartoli, A.: Fast explicit diffusion for accelerated features in nonlinear scale spaces. Trans. Pattern Anal. Mach. Intell. **34**(7), 1281–1298 (2011)
10. Csurka, G., Dance, C.R., Fan, L., Willamowski J., Bray, C.: Visual categorization with bags of keypoints. In: European Conference on Computer Vision (ECCV). Prague (2004)
11. Chih-Chung, C., Chih-Jen, L.: LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. (TIST) **2**(3), 27 (2011)
12. OpenCV Homepage. http://opencv.org/. Accessed 17 July 2017