AR-based Maintenance Verification Using Object State Recognition

Pavel Dvorak Radovan Josth Elisabetta Delponte

Konica Minolta Laboratory Europe
research.konicaminolta.eu

Abstract

In this work, we present a method for automatically recognizing an object's state with the objective of supporting a user to perform a set of tasks. The context of the application is that of Augmented Reality (AR) within the framework of industrial manufacturing and in particular for electronics rack maintenance. The method we propose makes use of the Histogram of Oriented Gradients (HOG) descriptor that has been widely used for different applications in Computer Vision. The test set on which this methodology has been evaluated comprises both artificially created and real world images and the results have demonstrated the suitability of the proposed algorithm in the target application domain. The novelty of our approach is that the system can automatically recognize the state of the object and provide immediate feedback to the operator using an AR visualization enabling fully automatic step-by-step instructions. The correct handling of the object can be automatically verified using the proposed system.

1. Introduction

Augmented Reality (AR), which uses visual search algorithms [1], is an emerging technology and one of the expected solutions for Industry 4.0 that represents the current trend of automation and digitization in manufacturing industry. The primary use case of such technology is supporting the workers' tasks that are not performed by them on a daily basis. Another significant group of target users is new employees. The main benefit from enabling the training of new employees to be faster and more efficient is related to the opportunity for cost savings. There is a large potential for the adoption of smart glasses, an emerging technology tightly connected to AR.

The problem of automatically recognizing objects from visual information has been extensively studied [2] [3]. However, in several applications concerned with monitoring of tasks in manufacturing and maintenance processes in industry achieving satisfactory results may be difficult due to the fact that different targeted objects may have highly similar appearances or may appear in variety of states that each need to be recognized. We propose a method that focuses on the task of automatically recognizing an object's state within the environment of an industrial manufacturing or maintenance process with the objective of verifying a set of tasks performed by an operator using Augmented Reality (AR).

By exploiting a cyber-physical system, in this case a pair of smart glasses that integrates a Holographic Optical Element (HOE) [4], the method we present is able to support the operator to successfully complete maintenance or manufacturing activities in an industrial context. The combination of the smart glasses and object state recognition method can inform the user in the case of incorrect handling of objects or other instruments and provide relevant information concerning the performed task. A further extension of this method includes aiding the navigation of the operator through the maintenance and manufacturing workflow to improve the process efficiency and reduce human errors. The system is able to provide feedback to the operator on how well he has performed the specified actions. Within the context of the manipulation of devices necessary to complete maintenance-type activities, accurate recognition of objects must be made quickly to provide the real-time support needed whilst performing the tasks. The benefit of such technology is an increased efficiency of the worker and a reduction in the number of errors caused by incorrect handling.

2. The method

The method presented is concerned with the recognition of similar objects (and focused only on 2D objects). It has been designed to distinguish different instances of an object that can have more than a single appearance or more than one state. An example of such a situation is depicted in Fig. 1. The first step in the object state recognition is the recognition of the object itself and for this purpose any suitable method providing the object's I.D. and its coordinates can be used. The proposed method makes use of a descriptor called Histogram of Oriented Gradients (HOG) [5], which was selected due to the robustness of the method to changes in light conditions, and its capability to recognize non-textured objects. HOG features provide information about the spatial and structural visual appearance of the object.

Reference Dataset. The reference dataset consists of images of each state of the object that is to be recognized, which are automatically aligned to the template picture, and a mask representing the differences between object states. Multiple states can be represented by a single mask. For each object, HOG features are computed for every single mask.

Fig. 1. Example of an object with three states represented by a single mask highlighted by the yellow rectangle. Sample object: a rack that needs to be populated with various types of electronic boards.

Object State Recognition. Technologies commonly employed in AR engines may be used for object detection such as the approach described by Girod et al. in [6]. After the object has been identified and its position determined, the image is transformed into the space of the template so that the areas to be verified in the query image are in the same position as their counterparts in the template image. The HOG features can then be computed from the target areas in the same manner as for reference images. The distance to each HOG descriptor from the particular area in reference images is computed and the object state with the minimum distance is selected as the suitable candidate. This calculation is simple for objects with a few states and simple masks, however, it may become increasingly computationally demanding for complicated objects.

3. Results and discussion

For the evaluation of the algorithm, two types of data were used: artificially created images and real-world images. The image created for the object state evaluation application comprises a set of pictures of objects of slightly varying appearance and with a limited number of states. Within this work, the success rate of object recognition within the dataset is presented in Tab. 1. The goal was to identify an object's state and therefore the dataset concentrated on 2D object images and did not deal with a large-scale visual search problem.

The results obtained with three different types of backgrounds (white, grey rectangles, and picture of electronic circuit) for 30 object states and with nine real-world object states (last column), whose samples are provided in Fig. 1 for several camera angles of view, indicate that the accuracy decreased due to the deformation caused by increasing the angle of view. These results also suggest that the performance achieved even at large angles of view may have sufficient accuracy depending on the type of usage. The experiments tested on artificial data were evaluated separately for all the data types that were used for testing the algorithm performance.

Tab. 1. Success rate of correct object state identification for three different types of backgrounds: white, grey rectangles, and a picture of electronic circuit; for 30 object states for several camera angles of view.

Angle of view / success rate (%)	Solid background	Rectangles background	Circuit background	Real-world images
50°	100	100	91	93
60°	100	100	86	88
70°	100	100	73	72
80°	100	96	50	46
85°	87	76	21	31
87°	47	40	8	23

The benefits of the algorithm developed in the present study lie in both its simplicity and speed as well as in its robustness to changes in light conditions and high accuracy with use-cases where the structural changes were distinguishable. The drawback of the proposed approach is the necessity of some manual work during the initialization process where masks are prepared. The automation of this part of the process is already under development in the authors' lab and part of future studies. The presented method has demonstrated promising results for the recognition of states of planar objects. Since current method can identify the state of 3D objects only if the angle of view is preserved, in future investigations, the authors will explore the extension of the algorithm proposed in this paper to the recognition of objects whose states are defined by 3D structural changes for varying camera angles of view.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 680633.

References

[1] X. Zhang, Z. Li, L. Zhang, W. Ma and H. Shum. Efficient indexing for large scale visual search. *ICCV*, 2009.

[2] C. M. Sukanya, G. Roopa and V. Paul. A Survey on Object Recognition Methods. *International Journal of Computer Science Engineering and Technology,* pp. 48-52, 2016.

[3] P. M. Roth and M. Winter. Survey of appearance-based Methods for object Recognition. Institute for Computer Graphics and Vision, Graz, 2008.

[4] Konica Minolta Inc. Wearable Communicator. 2017. http://www.konicaminolta.com/about/research/future/wcc

[5] N. Dalal and B. Triggs. Histograms of Oriented Gradients for Human Detection. CVPR. 2005.

[6] B. Girod, V. Chandrasekhar, R. Grzeszczuk, Y. A. Reznik. Mobile visual search: Architectures technologies and the emerging MPEG standard. IEEE Trans. Multimedia, vol. 18, no. 3, pp. 86-94, Mar. 2011.