Automatic visual inspection based on trajectory data

Georg Maier, Nico Mürdter, Robin Gruna,
Thomas Längle, and Jürgen Beyerer

Fraunhofer Institute of Optronics, System Technologies
and Image Exploitation IOSB
Fraunhoferstr. 1, 76131 Karlsruhe, Germany

Abstract Automatic inspection tasks have successfully been implemented in several industrial fields and are of growing importance. Visual inspection using optical sensors is wide spread due to the vast variety of different sensors, observable features and comparatively low prices. It seems obvious that corresponding systems are blind towards mechanical features and inspection of those typically requires highly specialized, inflexible and costly systems. Recently, we have shown in the context of sensor-based sorting that tracking objects over a time period allows deriving motion-based features which potentially enable discrimination of optically identical objects, although an optical sensor is used. In this paper, we take one step back from the specific application and study the classification of test objects based on their trajectories. The objects are observed while receiving a certain impulse. We further refrain from manually designing features but use raw coordinates as extracted from a series of images. The success of the method is demonstrated by discriminating spheres made of similar plastic types while bouncing off a plane.

Keywords: Machine vision, motion features, tracking.

1 Introduction

How can a cooked egg be distinguished from a raw one? Obviously, the difference cannot be determined by their appearance. A common household trick is to lay both eggs on a flat surface, rotate them like a spinning top and observe their rotation. While the boiled egg rotates uniformly, the raw egg performs a much more unstable movement due to the inertia of the liquid interior of the egg. This example demonstrates how an object to be tested is stimulated within the framework of an experiment in order to observe a characteristic behaviour. Due to their cognitive abilities, humans are immediately able to evaluate the observed movement behaviour and distinguish between the two objects. To a certain extent, an optical inspection is carried out here on the basis of a non-optical object property, namely the inertial tensor.

Away from the home kitchen, the task of sorting particles contained in a material stream according to certain criteria exists in several industrial fields. There exist two main types of systems for automation of the sorting process, namely mechanical sorting and sensor-based sorting [1]. Examples of mechanical sorting include sieving for separation based on size, sink–float processes for separating materials based on specific gravity and air-stream separation for separating particles with different aerodynamic characteristics. In sensor-based sorting, the characteristic used to distinguish particles from different classes determines the choice of the sensor used. For instance, RGB cameras are used for discrimination based on color, texture and shape, hyperspectral cameras can be used to retrieve information about the chemical composition of the particles and X-ray in order to measure the atomic density.

1.1 Problem statement and contribution

In the introductory example, we were interested in sorting objects according to mechanical properties. For many cases, this can be achieved by using mechanical sorting as discussed above. However, such processes typically suffer from a lack of flexibility, limited throughput and/or cost-intensive implementation. Sensor-based sorting systems appear to be an attractive alternative in all these regards. However, systems designed for high-throughput typically use imaging sensors and are hence by definition limited to optically perceivable characteristics.

In this paper, we propose a machine vision approach for the classification of objects based on non-optical properties in the context of automatic visual inspection. Our approach is based on the use of an area-scan camera in combination with object tracking methods. The classification is based on the trajectories of objects as observed in a specific scene. The movement of the objects is tracked using an image sequence recorded at a high temporal resolution. We present an experimental setup in which the objects are observed while receiving a certain impulse. The setup supports the generation of huge datasets by realizing a circulation of test objects. With the help of machine learning, it is shown that optically identical objects made of similar materials can be distinguished from each other based on their trajectories without any further feature engineering. The proposed approach can easily be extended for high throughput applications and requires inexpensive hardware.

1.2 Related work

Although the use of imaging sensors dominates in sensor-based sorting, other systems have been proposed to sort materials on the basis of non-optical properties. An example is performing classification based on impact resonant acoustic emissions. For instance, in [2], the applicability of such sorting systems is evaluated for the detection of damaged wheat kernels, including defects that are optically not perceivable. In [3], the authors propose a similar system for the sorting of End-of-Life vehicles' plastic materials. Their system further includes laser triangulation scanning to combine information regarding the size of single plastic flakes with features derived from the impact acoustic.

Several works have also discussed the idea of performing classification or quality assessment on the basis of motion information obtained from image data. For instance, in [4], two material properties of fabric, namely stiffness and area weight, are estimated based on motion induced by unknown wind forces. The authors propose a framework which includes extraction of the magnitudes of motion from video data, deriving statistical features and implements a regression model to estimate the material properties. In [5] a quality control system for application in an industrial setting based on the tracking of sputters during a laser-welding process is proposed. The events are tracked at a high frame rate in order to distinguish strong sputter events that are critical to the welding process from harmless ones. Regression of physical properties of objects from video data has also recently been studied. Motivated by gaining knowledge about how humans learn to predict motion of objects in the real world, the authors propose a model that allows predicting physical properties which are then fed into a physics engine in order to simulate the continuation of a dynamic scene [6].

Recently, we have proposed utilizing motion-based features for the characterization of materials in sensor-based sorting [7, 8]. We have shown that spheres made of different materials can be distinguished based on their motion while being transported on a conveyor belt. However, to that point, we restricted ourselves to using test objects made of strongly differing materials. Furthermore, rather primitive, hand-crafted features based on motion statistics were used and only passively induced interaction with the environment, in this case friction with the conveyor belt, was considered. The study presented here distinguishes itself from the former one in that very similar materials are used as test objects, no feature engineering is performed and an active impulse on the test objects is observed.

2 Materials and methods

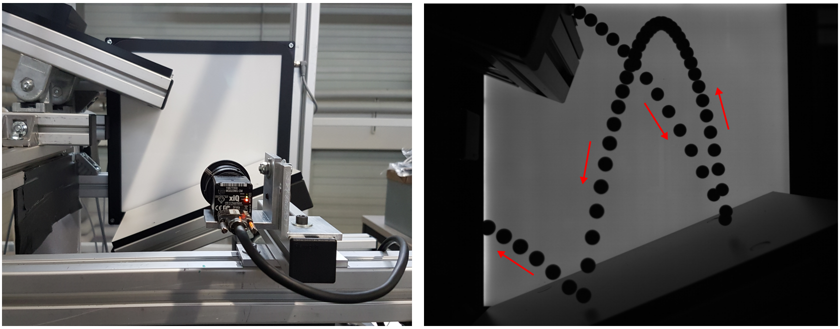
The following is a description of the setup designed to acquire a dataset and the methods used for analyzing the data.

2.1 Data acquisition

The phenomenon we want to observe in our experimental setup is elastic collision. We adopt the setup from [3] by using an inclined plane to accelerate the test objects and a second plane with which the objects collide, see Figure 9.1. Hence, we observe the test objects while bouncing off the second plane. After the collision(s), the test objects fall into a funnel and are re-applied on the inclined plane by using a Venturi loader.

With respect to the optical hardware, we use the camera Ximea xiQ MQ022 and an 8 mm lens. The camera is connected to a computer using the USB 3.0 interface. We further crop the image to an resolution of 1220×950 pixels and record images at 194 fps. Illumination is realized by using an LED back light.

As for test objects, we are interested in using objects of the same shape made of different, yet similar materials. For this purpose, we created 3D prints of sphere shaped test objects made of different plastics.



(a) Field of view of the camera.

(b) Exemplary motion pattern observed by the camera.

Figure 9.1: Impressions of the observed experimental situation.

We consider 8 materials from 4 different types of plastic, namely acrylonitrile butadiene styrene (ABS), polyamide (PA), polycarbonates (PC) and polypropylen (PP). All spheres have a diameter of 10 mm. An impression of the test objects is provided in Figure 9.2. It is important to note that the difference in appearance, i.e., color, is not used for the classification.

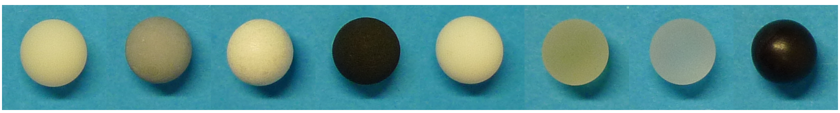


Figure 9.2: Photo of the test objects, from l. t. r.: ABS1, ABS2, PA1, PA2, PC1, PC2, PP1, PP2.

Our goal is to extract discrete time series data from the images which represent the path traveled by the test objects in 2D space. We neglect the third dimension as the objects move approximately only in one plane and the camera points perpendicular to that plane. In order to locate an object in an image, we apply image processing. In a first step, an image received from the camera is segmented using background subtraction using the implementation from the OpenCV library which is based on [9,10]. The resulting binary image is further pre-processed using morphological operations, namely erosion and subsequent dilation, followed by Gaussian filtering. In case an object is contained in the image, its contour is extracted using the implementation of [11] in OpenCV, yielding a measurement of form

$$p(t) := (x, y, t) \; . \tag{9.1}$$

The contextual attribute of the time series is a timestamp and the behavioural attribute is given by the 2D position of the center of the sphere in the image. A trajectory is then modelled as a set of subsequent measurements:

$$T := \{p(t_1), \ldots, p(t_n) \mid t_n \leq t_{n+1}\} \; . \tag{9.2}$$

As has been mentioned, our experimental setup enables circulation of a test object by re-applying it over and over again. Therefore, we need to determine which measurements belong to a single trajectory and which to different ones. We can group measurements to a single trajectory by determining the time difference between two measurements. At a constant recording speed, a trajectory is only valid if two consecutive points also originate from two directly consecutive images of the recording. A trajectory of a single pass can hence be formalized as

$$T := \{p(t_1), \ldots, p(t_n) \mid t_n < t_{n+1}, t_{n+1} - t_n \leq \epsilon\} \tag{9.3}$$

where $\epsilon := 1/\text{fps}$ is the time between two consecutive frames.

2.2 Data preparation and analysis

Prior to data analysis, we perform data cleaning in order to make sure we only work with trajectories without missing or possibly faulty measurements. For instance, we require that for each time point only a single measurement exists. The movement of each object is therefore described by a trajectory consisting of a temporally unambiguous point set to ensure that no faulty detection is included. We further only consider complete trajectories. The latter is ensured by exploiting a priori knowledge about the scene. We require that the first measurement lies within the area were the test objects enter the scene, i.e., the upper left corner with respect to Figure 9.1, and the last measurement where the objects leave the scene, i.e, the lower left corner. Using the described procedure, we create a dataset containing individual trajectories that are labeled with the corresponding material. For each of the 8 materials, the resulting dataset contains more than 10000 trajectories that were deemed valid.

We intend to use the recorded coordinates directly as the input for the classification without any further feature extraction. However, the trajectories are of varying length which results in a variable length of the feature vector, which is not supported by many classification algorithms. Therefore, for our experiments, we use two ways to extract trajectories of fixed length. The first method is extraction and padding. We calculate the median length of all recorded trajectories and use this length to either crop trajectories that contain more measurements or pad shorter trajectories to the length by filling up with zero-valued coordinates. The second method is based on geometric interpolation of the trajectories. The sampling is calculated with the help of a spline interpolation and the trajectories are up-sampled to 256 data points. We further discard the temporal component, i.e., the timestamp associated with each measurement, for the interpolation.

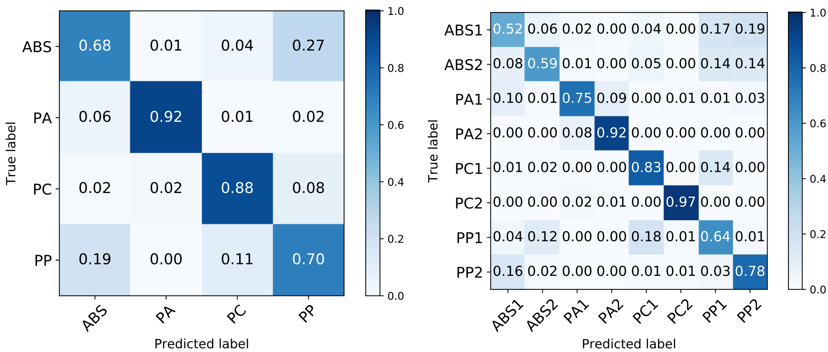
As a learning model, we use a support vector machine (SVM) with a radial basis function (RBF) kernel. The features, i.e., the coordinates, are standardized by removing the mean and scaling to unit variance.

3 Experimental results

For the experimental validation, we consider two types of classification problems. The first problem is to classify the material based on the trajectory data according to the plastic type as described in Section 2.1. For each type, there exist two individual test objects. For the second classification problem each individual test object is to be classified. We use 10000 trajectories of each material for the training and testing, resulting in a total of 80000 samples. The dataset is split into train and test sets whereas the size of the test set is 30% of the entire dataset.

Results when using cropped and padded trajectories are provided in Figure 9.3. As can be seen from Figure 9.3 (a), test objects of type PA and PC can be distinguished very well from the other types while a noteworthy amount of false classifications happen for ABS and PP.

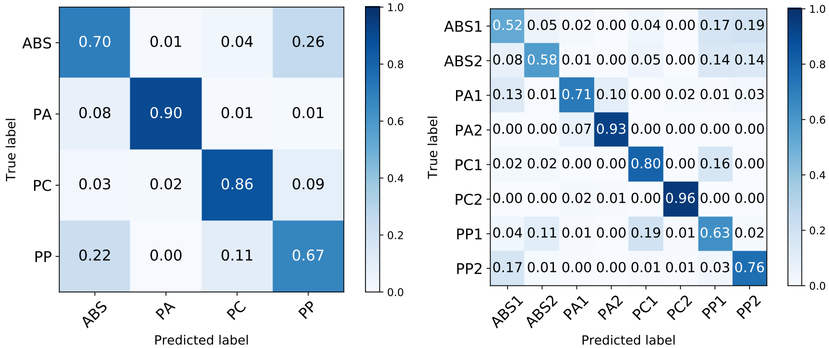
With respect to the latter, it can also be seen that the wrong classifications happen mainly between these two classes, i.e., ABS test objects are held falsely for PP and vice versa. From Figure 9.3 (b), the amount of false classifications with respect to the material within the plastic types can further be seen. It can be observed that with respect to PA, for instance, a large amount of false classification happen within the type, while this is not the case for PC. The false classifications happening between ABS and PP can be seen as well.



(a) Results using material types. (b) Results using individual test objects.

Figure 9.3: Normalized confusion matrices for classification results for cropped and padded trajectories.

Results for the geometrically interpolated trajectories are very similar to those using the cropped and padded trajectories, see Figure 9.4. Overall, a slight loss in classification performance can be observed, which might be explained by the loss of temporal information. More precisely, the data does not allow to extract information whether a test object passes the experiment faster than another one. In turn, results show that the cropping and padding did not harm the classification performance, which suggests that the additional step of interpolation is not necessary.



(a) Results using material types.

(b) Results using individual test objects.

Figure 9.4: Normalized confusion matrices for classification results for geometrically interpolated trajectories.

4 Conclusions

In this paper, we have extended our previous work on motion-based classification for application in automatic visual inspection. We designed an experimental setup suitable for the recording of huge amounts of data which makes application of purely data driven learning models feasible. Furthermore, we showed that modelling the trajectory as a time series with a timestamp as the contextual and 2D position as behavioural attribute suffices to achieve high classification performance, making feature engineering dispensable.

In the near future, we intend to gain more knowledge about the robustness of the classification performance by not only mixing materials but also shapes of the test objects. Furthermore, we want to integrate the approach in sensor-based sorting by incrementally adapting the observed situation. For instance, instead of an impulse induced by a collision, a rippled chute could be used as a transport mechanism. Lastly, the approach can be made applicable for high throughput applications by integrating multiobject tracking. This would allow motion-based classification of several objects observed at the same time.

Acknowledgment

This work was supported by the Fraunhofer Internal Programs under Grant No. Discover 614 068.

References

1. S. P. Gundupalli, S. Hait, and A. Thakur, “A review on automated sorting of source-separated municipal solid waste for recycling,” *Waste management*, vol. 60, pp. 56–74, 2017.
2. T. C. Pearson, A. E. Cetin, A. H. Tewfik, and R. P. Haff, “Feasibility of impact-acoustic emissions for detection of damaged wheat kernels,” *Digital Signal Processing*, vol. 17, no. 3, pp. 617–633, 2007.
3. J. Huang, C. Tian, J. Ren, and Z. Bian, “Study on Impact Acoustic—Visual Sensor-Based Sorting of ELV Plastic Materials,” *Sensors*, vol. 17, no. 6, p. 1325, 2017.
4. K. L. Bouman, B. Xiao, P. Battaglia, and W. T. Freeman, “Estimating the Material Properties of Fabric from Video,” in *Computer Vision (ICCV), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1984–1991.
5. M. Jager, S. Humbert, and F. A. Hamprecht, “Sputter Tracking for the Automatic Monitoring of Industrial Laser-Welding Processes,” *IEEE Transactions on Industrial Electronics*, vol. 55, no. 5, pp. 2177–2184, May 2008.
6. J. Wu, I. Yildirim, J. J. Lim, B. Freeman, and J. Tenenbaum, “Galileo: Perceiving physical object properties by integrating a physics engine with deep learning,” in *Advances in neural information processing systems*, 2015, pp. 127–135.
7. G. Maier, F. Pfaff, C. Pieper, R. Gruna, B. Noack, H. Kruggel-Emden, T. Längle, U. D. Hanebeck, S. Wirtz, and J. Scherer, Viktor Beyerer, “Improving material characterization in sensor-based sorting by utilizing motion information,” in *OCM 2017 - Optical Characterization of Materials*. KIT Scientific Publishing, 2017.
8. G. Maier, F. Pfaff, F. Becker, C. Pieper, R. Gruna, B. Noack, H. Kruggel-Emden, T. Längle, U. D. Hanebeck, S. Wirtz *et al.*, “Motion-based material characterization in sensor-based sorting,” *tm-Technisches Messen*, vol. 85, no. 3, pp. 202–210, 2018.
9. Z. Zivkovic, “Improved adaptive gaussian mixture model for background subtraction,” in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 2. IEEE, 2004, pp. 28–31.

10. Z. Zivkovic and F. Van Der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," *Pattern recognition letters*, vol. 27, no. 7, pp. 773–780, 2006.

11. S. Suzuki *et al.*, "Topological structural analysis of digitized binary images by border following," *Computer vision, graphics, and image processing*, vol. 30, no. 1, pp. 32–46, 1985.