Motion-based Visual Inspection of Optically Indiscernible Defects on the Example of Hazelnuts

Georg **Maier**[a], Anja **Shevchyk**[a], Merle **Flitter**[a], Robin **Gruna**[a], Thomas **Längle**[a], Uwe D. **Hanebeck**[b] and Jürgen **Beyerer**[a,c]

[a]*Fraunhofer IOSB, Karlsruhe, Germany, Institute of Optronics, System Technologies and Image Exploitation (e-mail: firstname.surname@iosb.fraunhofer.de)*

[b]*Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Karlsruhe Institute of Technology (KIT), Germany (e-mail: uwe.hanebeck@kit.edu)*

[c]*Vision and Fusion Laboratory (IES), Karlsruhe Institute of Technology (KIT), Germany (e-mail: juergen.beyerer@iosb.fraunhofer.de)*

ARTICLE INFO

Keywords:
Object trajectory
Motion classification
Sensor-based Sorting
Impact-acoustic

ABSTRACT

Automatic quality control has long been an integral part of the processing of food and agricultural products. Visual inspection offers solutions for many issues in this context and can be employed in the form of sensor-based sorting to automatically remove foreign and low quality entities from a product stream. However, these methods are limited to defects that can be made visible by the employed sensor, which usually restricts the system to defects appearing on the surface. An alternative non-visual solution lies in impact-acoustic methods, which do not suffer from this constraint. However, these are strongly limited in terms of material throughput and consequently not suitable for large scale industrial application. In this paper, we present a novel approach that performs inspection based on optically acquired motion data. A high-speed camera captures image sequences of test objects during a transportation process on a chute with a specific structured surface. The trajectory data is then used to classify test objects based on their motion behavior. The approach is evaluated experimentally on the example of distinguishing defect-free hazelnuts from ones that suffer from insect damage. Results show that by merely utilizing the motion data, a recognition rate of up to 80% for undamaged hazelnuts can be achieved. A major advantage of our approach is that it can be integrated in sensor-based sorting systems and is suitable for high throughput applications.

1. Introduction

Quality control has played a central role in the manufacturing and processing industry for many years. To humans, a variety of modalities is available to solve this task: test objects can be viewed, turned, felt, smelled, and so on. In many areas, performing quality control by human personnel is no longer feasible, because quantities are simply to large, employment of staff is uneconomical, and the process is highly error-prone and not reproducible [1]. This development gave rise to automatic quality control, which is vital to a vast variety of production and treatment processes. One of the most widespread kind of automatic quality control systems are visual inspection systems. The term *visual* does not imply a restriction to the range of the electromagnetic spectrum which is visible to humans, but rather incorporates imaging sensors. However, state-of-the-art machine vision systems are typically evaluating the product to be tested based on a static image. Compared with the above mentioned possibilities for humans, this is a severe limitation.

Ensuring product quality and safety is particularly crucial for the processing of food and agricultural products. Nuts are a comparatively expensive product and, due to the growing market in the area of healthy food, are enjoying increasing demand. Quality demands of consumers are particularly high. Although sensor-based sorting systems have successfully been employed to identify and sort out nuts with certain unwanted properties, for instance closed-shell pistachio nuts [2], those are, as mentioned above, by de-

sign limited in terms of the defects that can be detected. An alternative approach are so-called impact acoustic-based systems. Instead of imaging sensors, microphones are employed to analyze the acoustic emission of the product to be inspected during an impact with another rigid body. Various works have successfully demonstrated how such systems can be used for the detection of internal damages by analyzing the acoustic signal. However, a major drawback of the approach is its lack of scalability in terms of throughput, which in turn is a major economic concern. This is mainly because an individual sensor, i.e., microphone, is needed for each object to be tested simultaneously. Most likely for this reason, impact acoustic sorting systems appear to be rather of scientific interest than of industrial.

In other fields of study, evaluating quality on the basis of motion parameters is a fairly established means. Data sources for such approaches are for instance GPS data for the classification of vehicles in urban traffic [3] or absolute orientation sensors [4], but also include cases where information about the motion is derived from image data, for instance in computer-assisted sperm analysis [5], the identification of bird species [6] or reasoning about entities in video surveillance footage [7]. In [8], utilizing knowledge on the motion of particles during transportation in sensor-based sorting has been proposed.

In this paper, we present a novel approach to test agricultural products for visually indiscernible defects, such as interior damages, by means of visual inspection. A schematic illustration of the approach is provided in Fig. 1. The hypothesis of this study is that forcing test objects into an un-
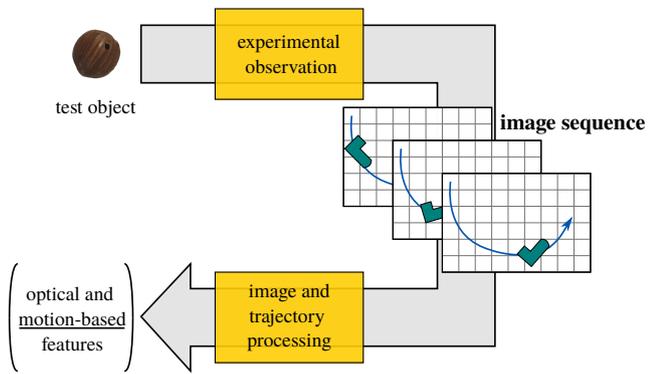
Figure 1: Schematic illustration of the proposed approach for motion-based visual inspection.

Figure 2: Schematic illustration of a chute-type sensor-based sorting systems. The turquoise objects represent particles to be accepted, typically the product, and the red ones those to be removed, for instance foreign particles. The field of view by the sensor is denoted by the yellow ray. The blue ray represents compressed air which is released to deflect an rejected object.

steady motion path can be used to reveal features influenced by the mechanical properties and allows reasoning about the test object. Our goal is to evaluate whether this effect can be exploited in machine vision tasks. The approach is based on the evaluation of the motion of test objects while facing "excitations" during transportation. Using machine learning, the motion data is used to detect faulty objects. Additionally to the motion data, the approach can naturally be combined with classical visual features, such as color, texture or shape. The advantage of the proposed approach is that, due to the usage of an imaging sensor, a huge number of objects can be tested simultaneously, allowing for high throughput applications. Furthermore, the concept integrates well into sensor-based sorting systems already established in industry. We demonstrate the success of our methods by designing and implementing an experimental setup and data processing pipeline for a case study with hazelnuts. The evaluated task involves the discrimination of hazelnuts suffering from insect damaged and undamaged ones by means of the introduced approach.

This paper is organized as follows. Following this brief introduction, related work is reviewed in Section 2. The data acquisition and processing as well as the experimental setup are introduced in Section 3. Experimental results for the detection of damaged hazelnuts are presented in Section 4. Finally, Section 5 concludes the paper.

2. Related Work

In many cases, sensor-based sorting systems can be employed for the inspection and physical cleaning of products by means of removing low-quality and foreign, potentially harmful, entities from a given product stream. The main fields of application of sensor-based sorting are recycling, for instance the preparation of glass [9] by removing materials harmful to the melting process such as stones and ceramic glass [10], mining, mainly to remove unwanted gangue from ore, e.g., copper-gold ore [11], as well as agricultural products and foodstuff. Regarding the latter, examples of applications are diverse, including the removal of fungus-infected wheat kernels [12], low-quality rice grains [13], and sunflower seeds [14], or quality insurance in bulgur
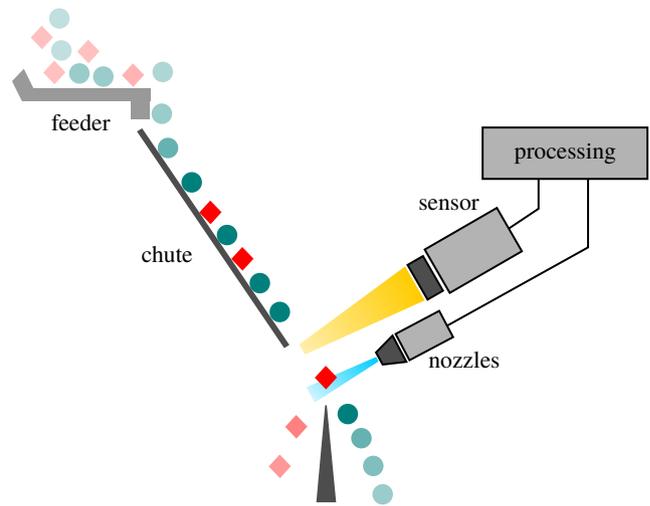
production [15]. A schematic illustration of a sensor-based sorting system is provided in Fig. 2. A material stream is fed into the system where it is further transported, for instance by means of a chute. The width of the chute and the bulk density are decisive for the achievable material throughput. Typically after being discharged, the material is perceived by one or multiple sensors. These commonly are scanning, imaging sensors, such as color or near-infrared cameras. Using image processing, individual objects are classified into an *accept* or *reject*[1] fraction based on a single observation. The sorting decision is executed by deflecting individual objects from the product flow. State-of-the-art systems use compressed-air nozzles for this purpose [16]. A major advantage of sensor-based sorting is that high throughput can be achieved, typically in the magnitude of multiple tonnes per hour. This is because objects distributed throughout the sorting width can be perceived by the sensor simultaneously and, thanks to the availability of fast sensors, can be transported at speeds typically ranging between $1\,\mathrm{m\,s^{-1}}$ and $10\,\mathrm{m\,s^{-1}}$. However, such imaging systems are, with the exceptions described below, limited to the detection of defects appearing on the surface or related to the geometry of the product.

There exist several works addressing the problem of designing automatic visual, i.e., imaging sensor based, inspection systems for the detection of visually indiscernible defects. In the context of sensor-based sorting, there are mainly two types of sensors which are employed for such tasks. The first type is Xray-based. Imaging via Xray Fluorescence (XRF) or Dual-Energy Xray transmission (DE-

---

[1]For the sake of completeness it is mentioned that theoretically, an arbitrary number of classes can be distinguished and separated. In industrial application, however, mostly binary sorting tasks are found due to complex mechanical handling for multi-way sorting.

XRT) is mainly found in the field of mining, for instance in order to assess the copper content in ores [17] or the beneficiation of coal [18]. However, applying Xray-based technology always comes at the cost – and this term is to be taken literally – of the necessity to satisfy rather strict safety regulations. The second type of imaging sensor that enables revealing information beyond the surface of a test object is Terahertz (Thz)-based imaging, as for instance presented in [19] in the context of black plastic sorting, which is particularly challenging in sensor-based sorting. However, this presented sensor system can still be regarded as being at an early development stage and further advances in sensor technology are required.

For foodstuff, in particular nuts and similar products, many approaches are based on impact acoustic. The main idea of these approaches is to record acoustic emissions caused by an entity of the product to be tested colliding with a rigid body. A schematic illustration of a typical setup is provided in Fig. 3. As can be seen, one or multiple microphones are used to record acoustic emissions during the collision with an impact plate. A possible extension of the setup lies in the inclusion of laser triggering systems to start the audio recording [20] or 3D cameras to also account for the size of the test objects during classification [21]. The choice of the best suitable microphone [22] or the material of the impact plate [22] depends on the detection task at hand. Means of processing the audio data are diverse and range from Fast Fourier Transform (FFT), Principle Component Analysis (PCA) to Deep Learning models [23, 24]. Examples of detection tasks include the discrimination of different genotypes of walnuts [23], classification of chestnuts according to their moisture level [20], detection of damaged rice [22] and wheat [25, 26] kernels, distinguishing almonds according to their thickness and hardness [27], and the detection of damaged [28] and hollow hazelnuts [29]. Besides the application in sorting and grading of agricultural products, impact acoustic is also used in the field of recycling, for instance for the discrimination of different plastic materials from automotive shredder residues [21] and sorting of municipal solid waste [24].

In the works mentioned above from the field of impact acoustic high recognition rates are reported, most of which typically above 90%. However, the throughput rates achieved are low and a serious obstacle for the industrial implementation. For instance, the system introduced in [28] is capable of the inspection of approximately 40 hazelnut kernels per second and its successor [30] of only 4 nuts per second. For the system presented in [26] a maximum throughput of 33 wheat kernels per second is reported. Hence, it can be concluded that material throughput is far from the quantities as for instance common in sensor-based sorting, where thousands of objects are treated per second.

The approach proposed in this paper is based on performing the detection of visually undetectable defects on the basis of motion perceived in image data. There exist several works that exploit visual motion information in order to inspect different products. Examples include reasoning about
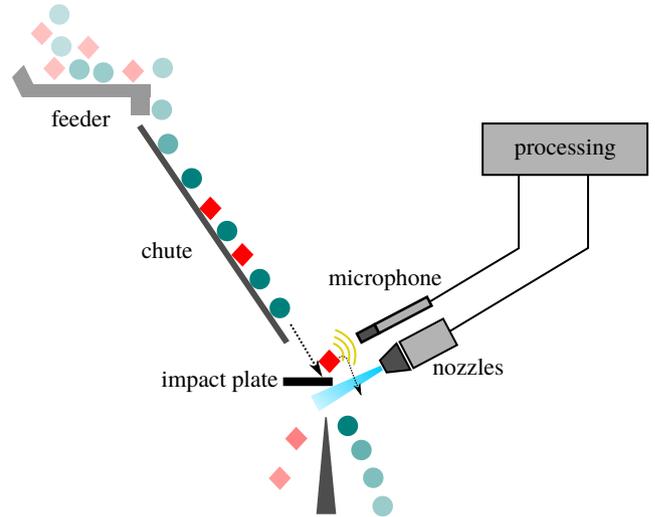
Figure 3: Schematic illustration of a basic impact acoustic setup featuring an impact plate and a microphone. As in Fig. 2, turquoise objects represent particles to be accepted and the red ones those to be removed.

the stiffness and area weight of fabric which is moved by unknown wind forces [31], the assessment of sputter events in a welding process [32], and also the general prediction of physical properties [33]. Recently, the approach of performing motion based classification has also been introduced in the context of sensor-based sorting [8]. Yet, this work only considers well-defined, artificial test objects. Also, an impact acoustic inspired setup is presented in [34] to record the motion of test objects while colliding with a plate. However, this also comes with the aforementioned drawbacks in terms of throughput, since, in contrast to the approach presented in this paper, it is hardly suited for observing multiple test objects simultaneously.

3. Methods and Materials

In the following, we introduce the experimental setup designed for this study. We further provide details on how experimental image data was acquired, the extraction of object trajectories from the image data, feature extraction, as well as the classification algorithms used.

3.1. Experimental Setups

Fig. 4 shows a free body diagram to visualize the forces and moments applied to a test object on a chute. For a basic analysis, we consider the 2D movement of a sphere, which approximates the shape of our test objects. The motion is determined by the contact angle of the cute and the body, the friction factor μ, the body mass m and the bodies mass moment of inertia J defined as

$$J = \int_V \vec{r}_\perp^2 \rho(\vec{r})\, dV \,, \tag{1}$$

where $\vec{r}$ denotes the distance of a given point of the test object to the rotation axis. The conservation laws are then given
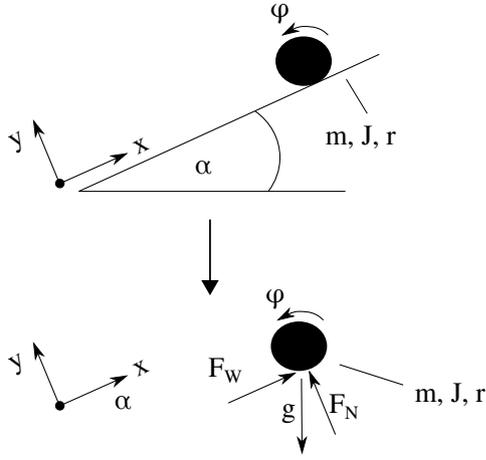
Figure 4: Free body diagram of a test object on a chute.

(a) Close-up of the structure. (b) Medium shot of the mat.

Figure 5: Photos of the polymer mat considered in the remainder to create a structured surface on the chute.

by

$$J\ddot{\varphi} = -F_W \cdot r$$
$$= -\mu \cdot m \cdot g \cdot \cos(\alpha) \cdot r \tag{2}$$

and

$$m \cdot \ddot{x} = -m \cdot g \cdot \sin(\alpha) + F_W$$
$$= -m \cdot g \cdot \sin(\alpha) + \mu \cdot m \cdot g \cdot \cos(\alpha) . \tag{3}$$

Boundary conditions and the kinematic relationship of the angular velocity $\dot{\varphi}$ and the vertical velocity $\dot{x}$ and $\dot{y}$ according to Fig. 4 are given by

$$\dot{x} = r \cdot \dot{\varphi} ,$$
$$\dot{y} = 0 . \tag{4}$$

Applied to our experiment, a geometry not symmetric to a central point generates a more complex trajectory with a 3D movement. We can assume an inhomogeneous density $\rho(\vec{r})$ for damaged hazelnuts. Accordingly, this results in differences with respect to the center of mass and the mass moment of inertia $\boldsymbol{J}$ compared with undamaged hazelnuts. This leads to an expected difference in motion in $\vec{x}$ and $\vec{\varphi}$. To amplify this effect, the experimental setup to be designed must ensure that test objects are forced into an unsteady trajectory. We further aim for a setup that integrates well into existing sensor-based sorting systems and supports the inspection of a high number of objects simultaneously. As a basis, we employ a chute serving the purpose of material transportation. For our experiments, this chute is inclined by $\alpha = 17.6°$ horizontally and has a width of 39.4 cm. A special feature of our chute is that it can be covered with a structured surface as shown in Fig. 5. This structured underground, which is provided by a polymer mat in our study, serves as a handicap for the motion of the test objects, forcing them into an unsteady motion path due to different impact angles.

After sliding down the chute, the test objects drop into a cone and further on into a flexible pipe. A venturi feeder is connected to the pipe and serves as a pneumatic conveyor. This way, the test objects are pushed by compressed air back on top of the chute. The compressed air is set to a constant value, so the objects are accelerated in the same way. Hence, the setup enables automatic circulation of the test objects, which facilitates data acquisition. A schematic illustration and a photo of the resulting experimental setup are provided in Fig. 6.

We mount a high-speed camera on the experimental setup such that an area of approximately 30.8 cm × 31.8 cm can be observed. The acquired images have a resolution of 980 px × 1016 px, resulting in a spatial resolution of approx. 0.31 mm px^{-1}. The camera model used in the experiments is Ximea xiQ MQ022. It is equipped with an 8 mm lens. The camera is connected to a computer using the USB 3.0 interface and enables recording of images at 181 fps in this setup. A LED flat dome is used to illuminate the viewing area of the camera.

3.2. Data Acquisition and Image Processing

Since continuously storing all images recorded by the camera as described in Section 3.1 would result in approximately 450 MB s^{-1} and is considered infeasible, we implement an online image processing and data reduction system. A schematic illustration of the software is provided in Fig. 7 and introduced in detail in the following. Our goal is to extract the image coordinates of the centroids of objects detected in each frame in order to create discrete time series data. The 2D position, the behavioral attribute of the discrete time series, is then stored together with a timestamp, which serves as the contextual attribute. Using the camera vendors C++ Application Programming Interface (API), we develop a software to set specific camera parameters, e.g., the integration time, and store the retrieved images in memory. For the pre-processing stage, we implement a threshold-based method to convert the color image to a binary mask image, where one value encodes the background and the other the foreground, i.e., the test object. The resulting binary image is further pre-processed using morphological operations, namely erosion and subsequent dilation, followed by Gaussian filtering. In order to be able to handle the high data throughput, these steps are executed on a graphics process-
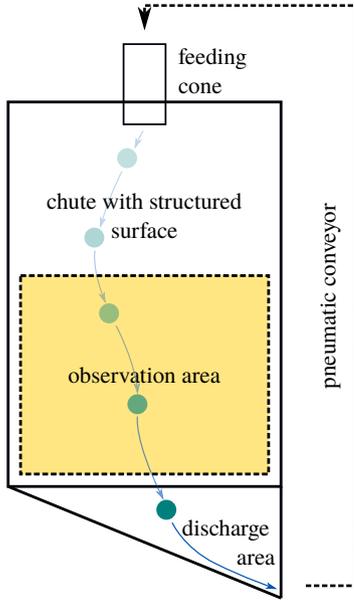
(a) Schematic illustration of the experimental setup.
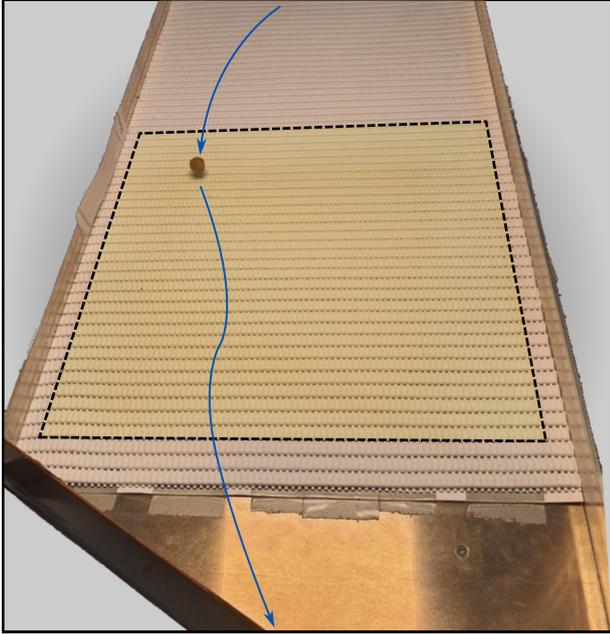
(b) Photo of the resulting experimental setup in the laboratory with the surface as shown in Fig. 5 mounted on the chute.

Figure 6: Illustrations of the designed experimental setup. The yellow, dashed square corresponds to the viewing area of the high speed camera. The blue path indicates a potential trajectory followed by a test object.

ing unit (GPU). Next, the contour of the object is extracted and the centroid is calculated during the detection stage. The centroid coordinates and the number of the frame it was detected in are eventually stored in a file, resulting in discrete measurements in the form of

$$p_t := (x_t, y_t) . \tag{5}$$

The timestamp t corresponds to the frame number the measurement stems from. To increase performance and exploit
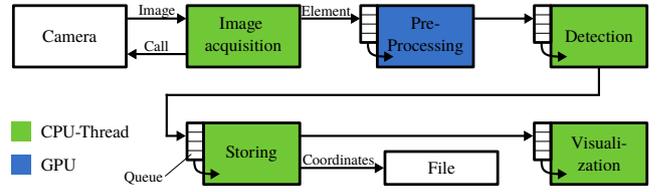
Figure 7: Schematic illustration of the image processing pipeline.

the advantages of multi-CPU systems, the software is designed in a pipeline structure such that the steps depicted in Fig. 7 can be run concurrently.

Up to this point, there merely exist data in form of unrelated, single point measurements. However, after having recorded the data, we want to model a trajectory as

$$T := \{p_{t_1}, \dots, p_{t_n} \mid t_n \leq t_{n+1}\} , \tag{6}$$

i.e., as a set of subsequent measurements. Using the contextual attribute of our data, i.e., the timestamp, we identify those measurements which stem from the same sample. Here, a sample is defined as a single run of a test object down the chute. By restricting ourselves to recording data for a single object at a time, i.e., there never occur two test objects in a single frame, we can identify the measurements stemming from a single pass by requiring that

$$T := \{p_{t_1}, \dots, p_{t_n} \mid t_n < t_{n+1}, t_{n+1} - t_n \leq \epsilon\} \tag{7}$$

where $\epsilon := 1/\text{fps}$ is the time between two consecutive frames. In other words, we exploit the knowledge that there is always at least one image without detection between two runs. It is important to note that although the test objects are fed individually into the system and our data acquisition system considers only one test object at a time, the system can be extended, for instance by integrating a real-time multiobject tracking system for sensor-based sorting as presented in [35], to handle multiple test objects occurring in the same frame. We further require that the number of points associated with an sample, i.e., $|T|$, is greater than a certain threshold which is based on the average number of points associated to the extracted samples. An exemplary visualization of resulting trajectories is shown in Fig. 8.

3.3. Trajectory Data Processing

In the following, we describe the implemented data processing pipeline with the goal of classifying individual trajectories. The implemented pipeline consists of four steps, namely pre-processing, feature extraction, post-processing and classification. During pre-processing, we determine the number of points we want to use per trajectory in order to obtain an initial feature vector of constant length. The number of points is determined as the mean length of all trajectories recorded for the considered classes. Trajectories containing more points are cropped, while trajectories containing less points are filled up with zero-valued coordinates. Furthermore, for each trajectory, we let

$$p_{t_s} := p_{t_1} \tag{8}$$
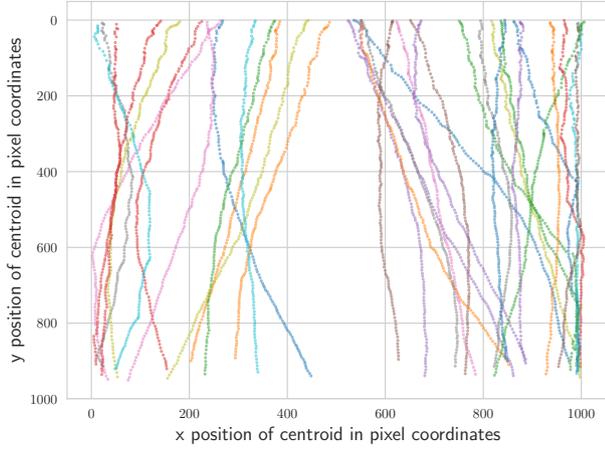
Figure 8: Exemplary visualization of acquired trajectories of samples from an undamaged hazelnut. The main running direction is from top to bottom. The different colors represent the different samples and the individual markers the corresponding centroids as extracted by the image processing system.

and align it to the origin by translating all coordinates according to

$$T := \{p_{t_1} - p_{t_s}, p_{t_2} - p_{t_s}, \dots, p_{t_n} - p_{t_s}\}. \quad (9)$$

This transformation is valid due to the translational symmetry of the stationary surface and results in our processing becoming invariant to wherever the trajectory entered the observation area of the camera.

For feature extraction, we consider approaches that have been proposed in the literature for the classification of objects based on their trajectory, although in rather different fields of study. For reasons of comprehensibility, we introduce names for the different approaches, however those do not necessarily comply with the original source.

Our first type of feature extraction is based on the idea presented in [7]. The authors propose using the coefficient feature space of spatiotemporal function approximations to classify objects in video surveillance footage. As approximation functions, least squares polynomials, Chebyshev polynomials and discrete Fourier transform are considered. A parameter for this kind of feature extraction is the degree used for the approximation functions, which also determines the size of the resulting feature vector. In the remainder, we refer to this kind of feature extraction by "coefficient feature space (CFS)".

The second type of feature extraction is based on the descriptive statistical analysis proposed in [3], which is originally proposed in the context of classification of urban vehicles based on GPS data. The authors propose a set of statistical features in order to derive motion profiles. In our implementation, we consider velocity, acceleration and turning angles for the calculation of such profiles. These profiles are eventually decomposed based on sinuosity and a deviation index. For details of the algorithm, the interested reader is

kindly referred to [3]. A parameter for this kind of feature extraction is the so-called decomposition threshold of the algorithm. With reference to the framework originally proposed in [3], our feature vector includes the mean as well as the standard deviation of the length of the resulting segments of each decomposition class, the count of profile changes per decomposition class, as well as the ratio of segment lengths per decomposition class in relation to the overall profile. The resulting feature vector has a length of 45. Based on the name of the publication, we will call these features "Physics of Movement (POM) features" in the following.

Lastly, we propose a third type of feature extraction which is based on velocity and turbulence. We calculate velocities in a sliding window manner, where the window size w is the parameter for this kind of feature extraction. Velocities are then calculated for both directional components individually as given by

$$V_x := \{x_{t_w} - x_{t_1}, x_{t_{w+1}} - x_{t_2}, \dots, x_{t_n} - x_{t_{n-(w-1)}}\} \quad (10)$$

and for y likewise. Turbulence intensity (TI) is a feature typically used in scientific applications such as pipe flow simulations [36] or also for optimizing wind farm design [37]. For each time window, we calculate the average velocity $\bar{v}$ over the entire window as well as the velocities v_i between all time-wise consecutive coordinates within the window. Fluctuations in velocity are then calculated for both directional components individual, on the example of the x component according to

$$v_{f,x}^2 := \frac{1}{w-1} \sum_{i=1}^{w} (v_{i,x} - \bar{v}_x)^2 \quad (11)$$

and then used for the calculation of TI as given by

$$\mathrm{TI} := \frac{1}{U} \sqrt{\frac{1}{2}(v_{f,x}^2 + v_{f,y}^2)} \quad (12)$$

where U is the mean velocity over all known samples. The calculation for the directional y component follows likewise. The resulting feature vector includes $(n-w)$ velocity-related features for the x and y component and $(n-w-1)$ turbulence-related features for the x and y component, resulting in a total feature vector length of $(4n - 4w - 2)$. Since this type of feature extraction does not strongly relate to an approach known from the literature, we will refer to this kind of feature extraction as the proposed approach.

In the post-processing step, we split the data into a train and test set by the ratios 2/3 and 1/3, respectively. We further determine the mean and standard deviation of the features from the train set and then transform both the test and train set by removing the mean and scaling to unit variance.

For classification, we consider a support vector machine (SVM) with a radial basis function kernel.

4. Experimental Results

In our case study, we consider two classes of hazelnuts, namely undamaged hazelnuts and hazelnuts that suffer

(a) Photo of the hazelnut samples that suffered from insect damage. The whole drilled by a nut weevil can also be seen. The weight of the nuts are (FLTR) 0.98 g, 1.73 g and 1.59 g.

(b) Photo of the hazelnut samples without any damage. The weight of the nuts are (FLTR) 2.44 g, 1.55 g, 1.66 g and 2.39 g.

Figure 9: Photos of the a damaged and b undamaged hazelnuts.

from insect damage. Photos of the test objects are provided in Fig. 9. From Fig. 9a it can be seen that the insect damage was most likely caused by a beetle called *Curculio nucum* (nut weevil). Naturally, the different nuts further differ in shape, size and weight. For each of the nuts, we acquire 5000 valid samples, i.e., trajectories satisfying the conditions introduced in Section 3.2. The average count of data points in these trajectories, i.e., the mean length, is 66 and all trajectories are cropped or filled to this length. The total set of samples is then randomly split into the train and test set. Since our dataset includes 20000 samples of undamaged and 15000 samples of damaged hazelnut trajectories, we consider the corresponding classification baselines at approx. 57% for undamaged and 43% for damaged hazelnuts. This is the classification result that would be obtained when always choosing the corresponding class.

For the evaluation, we implement the data processing approaches presented in Section 3.3 and assess their suitability for the task of discriminating damaged hazelnuts from undamaged ones. For quantitative assessment, we determine the accuracy of the classification, which is defined as the ratio of correctly classified samples to all samples. As has been mentioned in Section 3.3, all three types of feature extraction considered require parameterization, whereas the optimal one is initially unknown. Namely the parameters are the degree of the approximation functions for the CFS approach, the decomposition threshold for the POM features and the window size for the proposed feature set, respectively. We determined the best parameterization in a brute force manner. As an example, the results for all plausible window sizes for the proposed feature set is shown in Fig. 10.

The respective best classification results of the individual methods are shown in Table 1. As can be seen, employing

Figure 10: Exemplary visualization of the process for obtaining the best parameterization per approach in a brute force manner, here on the example of the proposed feature extraction approach. By testing each plausible parameterization, the best one can be determined, in this case a window size of 5.

Table 1
Summary of the results for all considered types of feature extraction for the classification task of distinguishing damaged from undamaged hazelnuts. The respective best results are printed in bold.

Features	Accuracy undamaged	Accuracy damaged	Total accuracy
CFS: Least-squares polynomials	79 %	48 %	66 %
CFS: Chebyshev polynomials	79 %	48 %	66 %
CFS: Discrete Fourier transform	**83 %**	43 %	66 %
POM features	77 %	54 %	67 %
Proposed features	80 %	**70 %**	**75 %**

CFS features always yields a result of 66 % total accuracy and employing POM features performs fairly similar, resulting in 67 % total accuracy. In all cases, the accuracy for undamaged hazelnuts is well above the baseline, yet the accuracy for damaged hazelnuts only marginally. Best results are obtained by using the proposed feature set consisting of velocity and turbulence related features. In this case, accuracy for both classes, 80 % for undamaged and 70 % for damaged hazelnuts, is well above the baseline.

Based on these results, it can be concluded that analyzing the motion of the test object yields meaningful features to support the discrimination of different classes of test objects. While in this study these features were used as the only ones for classification, we expect a high potential of realizing a more robust classification by combining this new set of features with traditional features derived from image data, such as color, texture and/or shape. Regarding the example of hazelnuts at hand, we were also able to observe that the structured surface of the chute supports a rotation of the hazelnuts while going down the chute. Since several frames are taken of each nut by the high-speed camera, this also allows optical inspection from several sides without the need for a multi-camera system. In the case of the assessed insect

damage, this for instance increases the chances of spotting the hole. Last but not least, the main advantage of the proposed approach is that it is designed in such a way that it can directly be integrated into sensor-based sorting and enables quality inspection of thousands of objects per second. This way, high throughput can be achieved, making it attractive for industrial application.

5. Conclusion

We have presented a novel approach for inspection and sorting tasks where defects that cannot be detected using static image data, for instance internal defects, can be identified. The approach is based on a high speed camera and real-time image processing to extract the position of individual test objects throughout their trajectory. This motion data is then used as basic features to classify the different objects. For this purpose, we considered different feature extraction methods from the literature and proposed the adaption of an approach well-known from other fields of study. We evaluated the approach on the inspection task of identifying hazelnuts that suffer from insect damage. Our results show that a discrimination of damaged and undamaged hazelnuts is possible at an accuracy of 80% for undamaged hazelnuts. The main advantage of the proposed approach is that it can be integrated into sensor-based sorting systems and is able to handle a large number of test objects simultaneously, which in turn results in high throughput and economic benefits.

In the future, we are interested in running experiments for tasks from different application domains, such as discriminating different polymer types for recycling applications. Also, we are particularly interested in experiments with densely distributed test objects in the images. In such situations, collisions of multiple test object can occur, which may reveal further interesting features to reason about the objects. Also, we expect that omitting feature engineering and employing purely data driven methods, for instance deep learning, may result in a boost in classification accuracy.

References

[1] J. Beyerer, F. León, C. Frese, Machine Vision: Automated Visual Inspection: Theory, Practice and Applications, Springer Berlin Heidelberg, 2015.

[2] T. Pearson, N. Toyofuku, Automated sorting of pistachio nuts with closed shells, Applied Engineering in Agriculture 16 (2000) 91.

[3] S. Dodge, R. Weibel, E. Forootan, Revealing the physics of movement: Comparing the similarity of movement characteristics of different types of moving objects, Computers, Environment and Urban Systems 33 (2009) 419–434.

[4] J. P. McNamara, C. Borden, Observations on the Movement of Coarse Gravel using Implanted Motion-sensing Radio Transmitters, Hydrological Processes 18 (2004) 1871–1884.

[5] D. Kime, K. Van Look, B. McAllister, G. Huyskens, E. Rurangwa, F. Ollevier, Computer-assisted sperm analysis (CASA) as a tool for monitoring sperm quality in fish, Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology 130 (2001) 425–433.

[6] J. Atanbori, W. Duan, E. Shaw, K. Appiah, P. Dickinson, Classification of bird species from video using appearance and motion features, Ecological Informatics 48 (2018) 12–23.

[7] A. Naftel, S. Khalid, Classifying spatiotemporal object trajectories using unsupervised learning in the coefficient feature space, Multimedia Systems 12 (2006) 227–238.

[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 85 (2018) 202–210.

[9] N. Dias, I. Garrinhas, A. Maximo, N. Belo, P. Roque, M. T. Carvalho, Recovery of glass from the inert fraction refused by MBT plants in a pilot plant, Waste Management 46 (2015) 201–211.

[10] G. Bonifazi, S. Serranti, Imaging spectroscopy based strategies for ceramic glass contaminants removal in glass recycling, Waste Management 26 (2006) 627–639.

[11] T. Phiri, H. J. Glass, P. Mwamba, Development of a strategy and interpretation of the NIR spectra for application in automated sorting, Minerals Engineering 127 (2018) 224–231.

[12] S. R. Delwiche, T. C. Pearson, D. L. Brabec, High-Speed Optical Sorting of Soft Wheat for Reduction of Deoxynivalenol, Plant Disease 89 (2005) 1214–1219.

[13] C.-Y. Wee, R. Paramesran, F. Takeda, Sorting of rice grains using Zernike moments, Journal of Real-Time Image Processing 4 (2009) 353.

[14] Z. Luan, C. Li, S. Ding, M. Wei, Y. Yang, Sunflower seed sorting based on Convolutional Neural Network, in: Eleventh International Conference on Graphics and Image Processing (ICGIP 2019), volume 11373, International Society for Optics and Photonics, 2020, p. 113731K. URL: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11373/113731K/Sunflower-seed-sorting-based-on-Convolutional-Neural-Network/10.1117/12.2557789.short. doi:10.1117/12.2557789.

[15] M. Bayram, M. D. Öner, Determination of applicability and effects of colour sorting system in bulgur production line, Journal of Food Engineering 74 (2006) 232–239.

[16] C. Robben, H. Wotruba, Sensor-Based Ore Sorting Technology in Mining—Past, Present and Future, Minerals 9 (2019) 523.

[17] S. Nadolski, M. Samuels, B. Klein, C. J. Hart, Evaluation of bulk and particle sensor-based sorting systems for the New Afton block caving operation, Minerals Engineering 121 (2018) 169–179.

[18] L. von Ketelhodt, C. Bergmann, Dual energy X-ray transmission sorting of coal, Journal of the Southern African Institute of Mining and Metallurgy 110 (2010) 371–378.

[19] A. Küter, S. Reible, T. Geibig, D. Nüßler, N. Pohl, THz imaging for recycling of black plastics, tm - Technisches Messen 85 (2017) 191–201.

[20] F. Kurtulmus, S. Öztüfekci, I. Kavdir, Classification of chestnuts according to moisture levels using impact sound analysis and machine learning, Journal of Food Measurement and Characterization 12 (2018) 2819–2834.

[21] J. Huang, C. Tian, J. Ren, Z. Bian, Study on Impact Acoustic—Visual Sensor-Based Sorting of ELV Plastic Materials, Sensors 17 (2017) 1325.

[22] J. Buerano, J. Zalameda, R. Ruiz, Microphone system optimization for free fall impact acoustic method of detection of rice kernel damage, Computers and Electronics in Agriculture 85 (2012) 140–148.

[23] K. Simin, A. Hosainpour, M. Asghar, Detection of walnut varieties using impact acoustics and artificial neural networks (anns), Modern Applied Science 6 (2012).

[24] G. Lu, Y. Wang, H. Yang, J. Zou, One-dimensional convolutional neural networks for acoustic waste sorting, Journal of Cleaner Production (2020) 122393.

[25] T. Pearson, A. Cetinb, A. Tewfikc, R. Haff, Feasibility of impact-acoustic emissions for detection of damaged wheat kernels, Digital Signal Processing 17 (2007) 617–633.

[26] M. Guo, Y. Ma, X. Yang, R. W. Mankin, Detection of damaged wheat kernels using an impact acoustic signal processing technique based on gaussian modelling and an improved extreme learning machine algorithm, biosystems engineering 184 (2019) 37–44.

[27] A. Reshadsedghi, A. Mahmoudi, et al., Detection of almond varieties

using impact acoustics and artificial neural networks., International Journal of Agriculture and Crop Sciences (IJACS) 6 (2013) 1008–1017.

[28] I. Onaran, B. Dulek, T. Pearson, Y. Yardimci, A. Cetin, Detection of empty hazelnuts from fully developed nuts by impact acoustics, 13th European Signal Processing Conference (2015) 1–4.

[29] M. Farhadi, Y. Abbaspour-Gilandeh, A. Mahmoudi, J. M. Maja, An integrated system of artificial intelligence and signal processing techniques for the sorting and grading of nuts, Applied Sciences 10 (2020) 3315.

[30] A. A. Enis Cetin, T. Pearson, A. Sevimli, System for removing shell pieces from hazelnut kernels using impact vibration analysis, Computers and Electronics in Agriculture 101 (2014) 11–16.

[31] K. L. Bouman, B. Xiao, P. Battaglia, 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.

[32] M. Jager, S. Humbert, F. A. Hamprecht, Sputter Tracking for the Automatic Monitoring of Industrial Laser-Welding Processes, IEEE Transactions on Industrial Electronics 55 (2008) 2177–2184.

[33] J. Wu, I. Yildirim, J. J. Lim, B. Freeman, 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.

[34] G. Maier, N. Murdter, R. Gruna, T. Langle, J. Beyerer, Automatic visual inspection based on trajectory data, in: OCM 2019-Optical Characterization of Materials: Conference Proceedings, KIT Scientific Publishing, 2019, p. 87.

[35] G. Maier, F. Pfaff, C. Pieper, R. Gruna, B. Noack, H. Kruggel-Emden, T. Längle, U. D. Hanebeck, S. Wirtz, V. Scherer, et al., Experimental evaluation of a novel sensor-based sorting approach featuring predictive real-time multiobject tracking, IEEE Transactions on Industrial Electronics (2020).

[36] F. Russo, N. T. Basse, Scaling of turbulence intensity for low-speed flow in smooth pipes, Flow Measurement and Instrumentation 52 (2016) 101–114.

[37] R. J. Barthelmie, S. T. Frandsen, M. Nielsen, S. Pryor, P.-E. Rethore, H. E. Jørgensen, Modelling and measurements of power losses and turbulence intensity in wind turbine wakes at middelgrunden offshore wind farm, Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology 10 (2007) 517–528.