3-D Refuse-Derived Fuel Particle Tracking-by-Detection Using a Plenoptic Camera System

Miao Zhang, Markus Vogelbacher, Veit Hagenmeyer, Krasimir Aleksandrov, Hans-Joachim Gehrmann, and Jörg Matthes

***Abstract*— Multiple-particle tracking-by-detection is a widely investigated issue in image processing. This article presents approaches to detecting and tracking various refuse-derived fuel particles in an industrial environment using a plenoptic camera system, which is able to yield 2-D gray value information and 3-D point clouds with noticeable fluctuations. The presented approaches, including an innovative combined detection method and a postprocessing framework for multiple-particle tracking, aim at making the most of the acquired 2-D and 3-D information to deal with the fluctuations of the measuring system. The proposed novel detection method fuses the captured 2-D gray value information and 3-D point clouds, which is superior to applying single information. Subsequently, the particles are tracked by the linear Kalman filter, the 2.5-D global nearest neighbor (GNN), and the joint probabilistic data association (JPDA) approach, respectively. As a result of several inaccurate detection results caused by the measuring system, the initial tracking results contain faulty and incomplete tracklets that entail a postprocessing process. The developed postprocessing approach based merely on particle motion similarity benefits a precise tracking performance by eliminating faulty tracklets, deleting outliers, connecting tracklets, and fusing trajectories. The proposed approaches are quantitatively assessed with manually labeled ground-truth datasets to prove their availability and adequacy as well. The presented combined detection method provides the highest F_1-score, and the proposed postprocessing framework enhances the tracking performance significantly with regard to several recommended evaluation indices.**

***Index Terms*— Multiple-particle detection, multiple-particle tracking, plenoptic camera, tracking-by-detection, tracklet connection.**

I. Introduction

REFUSE-DERIVED fuel (RDF) has been widely utilized in various industrial processes, such as firing in the cement industry as a heat source. RDF impacts positively on the CO_2 balance of combustion processes owing to its biogenic proportion. In addition, RDF is cost-efficient compared to other conventional fuels. Nevertheless, the aerodynamic properties of RDF can be relatively complicated and unsteady because of the complex composition, which increases the difficulty of a controllable and highly effective application of RDF. To address the problem, several researchers attempted to investigate the properties of RDFs [1], [2], [3]. In this article, we present novel camera-based approaches to detecting and tracking RDF particles spatially to determine RDF's flight and combustion properties, e.g., trajectories and velocities, as well as ignition and dwell time, for further analysis of flight and combustion properties.

The concept of tracking-by-detection consists of two major tasks. First, objects captured in the images are detected and localized. Second, tracking algorithms are applied to the detections to perform data association and, thus, form object trajectories afterward.

Multiple-object detection is an essential task in various applications, for instance, surveillance systems, robotics, and biomedicine. Plenty of algorithms are developed to detect diverse objects, such as human beings or cells. In our work, we focus on detecting small fuel particles. In general, such particles can be detected using gray value threshold-based methods, e.g., the OTSU threshold selection approach [4] and scale-invariant feature transform (SIFT) [5]. Although SIFT is initially a feature extraction method, it can be used directly for particle detection, as shown in [6]. Applications of the approaches to detecting fuel particles are presented in [7], where burning fuel particles inside a combustion chamber are detected by applying a simple threshold to estimated foreground models. Further practical usage of the detection algorithms is exampled in [8], where the authors propose a detection and tracking method of firing particles from an image sequence based on gray level change. Principally, the detection of microscopic objects, for instance, microscopic particles and cells, shares the analogous solutions to detecting fuel particles. In [9], a variety of methods to detect or classify particles in electron microscopy are reviewed, for instance, cross correlation-based approaches [10], texture-based methods [11], and artificial neural networks [12]. Later, Irshad et al. [13] outline nuclei detection approaches for modern medicine, including thresholding approaches [14],

Manuscript received 20 August 2022; revised 27 September 2022; accepted 17 October 2022. Date of publication 28 October 2022; date of current version 11 November 2022. This work was supported by the AiF—German Federation of Industrial Research Associations under Grant 20410N. The Associate Editor coordinating the review process was Dr. Huan Liu. *(Corresponding author: Miao Zhang.)*

Miao Zhang, Markus Vogelbacher, Veit Hagenmeyer, and Jörg Matthes are with the Institute of Automation and Applied Informatics, Karlsruhe Institute of Technology, 76344 Karlsruhe, Germany (e-mail: miao.zhang@kit.edu; markus.vogelbacher@kit.edu; veit.hagenmeyer@kit.edu; joerg.matthes@kit.edu).

Krasimir Aleksandrov and Hans-Joachim Gehrmann are with the Institute for Technical Chemistry, Karlsruhe Institute of Technology, 76344 Karlsruhe, Germany (e-mail: krasimir.aleksandrov@kit.edu; hans-joachim.gehrmann@kitedu).

Digital Object Identifier 10.1109/TIM.2022.3217858

active contour model [15], and K-means clustering [16]. Several methods, such as the thresholding and cross correlation approach, depend strongly on local gray value distribution that could vary significantly in actual industrial combustion processes. The artificial neural network might provide precise performance. Nevertheless, it entails relatively high computational costs and learning datasets. K-means clustering relies not only on pixel values but also on clusters' initialization, limiting its utilization.

In the case where scenes can be captured three-dimensionally, such as using stereo cameras, RGB-D cameras, ToF cameras, or plenoptic cameras, clustering methods can be applied to the obtained point clouds to implement object detection. A new idea using mean shift clustering candidate segmentation for people detection based on the point clouds gathered by an RGB-D camera is introduced in [17]. Chiang et al. [18] propose a fast clustering method on the basis of density-based spatial clustering of applications with noise (DBSCAN) algorithm [19] to realize traffic detection for self-driving technology. Huang et al. [20] recommend a novel center-based clustering algorithm quasi-cluster centers (QCCs) that distinguish dense neighbors from K-nearest neighbor (NN) and extend clusters regarding the defined QCCs and the corresponding dense neighbors. The results of the proposed algorithm on synthetic and practical datasets proved its adequacy and accuracy. Wang et al. [21] present a new clustering method based on morphological operations (MOs), which utilizes morphological dilation to connect the data points based on their adjacency. Experimental results indicate the robustness and universality of the proposed method for clustering 2-D or 3-D data.

In recent years, researchers attempted to fuse 2-D- and 3-D information of a particular scene to realize more possibilities in computer vision. Halima et al. [22] contribute to detecting and tracking human heads by fusing depth and 2-D thermal information. Attamimi et al. [23] propose an object detection approach based on a particle filter with integrated color, texture, and depth information for the detection of everyday items. Justen [24] conducts 3-D-image processing, such as 3-D-object reconstruction and object classification, based on gray value images and individual depth information, which is deemed a new 2-D/3-D camera concept. Ophoff et al. [25] demonstrate real-time object detection by fusion of standard RGB data with depth information to increase the performance of current detection networks. The majority of these fusion methods achieve sufficient accuracy in detecting relatively large objects. Applying the approaches to industrial burning fuel particles seems currently not feasible due to the lack of specific object information and features.

As mentioned, multiple-object detection is followed by multiple-object tracking, which enables the association of detections to acquire object trajectories. In general, object tracking consists of two major parts: a tracking filter and a data association approach. The tracking filter aims at predicting and updating particle positions in the time dimension according to the previous positions and the currently assigned detections, such as the Kalman filter [26] and particle filters [27]. Over the last several decades, data association algorithms have been developed and modified. Deterministic approaches, for instance, NN and global NN (GNN) [28], assign a single detection within the gate around a particular prediction to the corresponding track by considering local or global minimal assignment costs. The gate of a prediction is usually determined by a threshold of the Mahalanobis distance that considers the directional scatterings individually. Compared to deterministic approaches, probabilistic methods, such as probabilistic data association (PDA) [29] and joint PDA (JPDA) [30], weigh all detections within the gate to follow objects instead of assigning simply one detection and discarding the others. It is obvious that probabilistic algorithms increase the opportunity for correct data association in cluttered environments, while deterministic methods could assign noise to tracks that lead to entirely inaccurate results. Nevertheless, deterministic methods entail less computational costs in comparison with probabilistic methods since probabilistic methods take all detections within the gate into consideration and enlarge the gating area in accordance with the detections afterward. When training data is available, deep learning methods could provide an alternative solution. Milan et al. [31] present an approach to online multitarget tracking based on recurrent neural networks (RNNs) and apply the approach to a benchmark of people tracking. Ahmad et al. [32] investigate the performance of a faster region convolutional neural network (Faster-RCNN) in combination with Generic Object Tracking Using Regression Networks (GOTURN) architecture for human-being tracking. Deep learning methods could achieve sufficient accuracy, as demonstrated by the authors. Nevertheless, the computational cost is raised, and in addition, extensive training data and specific features of the tracked objects are necessary.

The above-described tracking approaches are feasible in both 2-D and 3-D. Notwithstanding, the recorded 3-D information ought to deliver satisfactorily steady depth information within a video sequence to ensure an accurate tracking performance. Unfortunately, the available measurement systems fail to fulfill the condition ordinarily when tracking small objects in complicated industrial processes, like small burning fuel particles. Due to the partial overexposure caused by burning particles and their small sizes, the measurement system is not capable of providing 3-D information with slight fluctuation. Under this circumstance, direct 3-D tracking is impracticable. However, the acquired depth information can be utilized implicitly, as presented in [33], where the depth information is considered a threshold factor. This new concept deals with relatively unsteady and inaccurate 3-D information in object tracking, for instance, 3-D information gathered in complex industrial processes. The 3-D data can be partly applied as a constraint for 2-D tracking to optimize the tracking performance.

Despite the efforts invested in tracking approaches, some tracking issues remain inevitable. These issues can be resulted in detection errors, e.g., false detections or duplications, or in the tracking algorithm itself. Hence, researchers contribute to proposing postprocessing approaches for multiple-object tracking. In [34] and [35], approaches to reconnecting tracklets by computing the apparent similarity of the repre-

sentative objects and the motion similarity of the tracklets are demonstrated. Nevertheless, for detecting small particles, such methods are not definitely adequate to amend trajectories of target objects that are too small and too similar to extract independent appearance features. Under this circumstance, novel approaches relying principally on motion similarity to processing tracking results ought to be developed.

In this article, we introduce and demonstrate methods to process images captured by a plenoptic camera for detecting and tracking RDF particles spatially. In order to enhance the accuracy of multiple-particle detection, we propose a novel combined detection method that combines the detection results of two distinct approaches and takes not only the 2-D gray value information but also 3-D point clouds into account for detection, which we first introduced in [36]. This proposed combined approach compensates for the deficiencies of the two distinct methods and achieves the highest F_1-score. Subsequently, the detections are linked to acquire 2-D particle tracklets using a linear Kalman filter together with a modified data association algorithm that uses additional 3-D information for gating. Afterward, we present a postprocessing framework to deal with issues that occurred in the tracklets and connect tracklets into complete 2-D particle trajectories. The developed postprocessing approach first eliminates faulty tracklets in accordance with the length and motion properties (velocity and regularity) of the tracklet. Based on the motion properties, outliers in each tracklet can be deleted. Subsequently, the remaining tracklets are connected or merged by computing related paring costs, which take motion properties and time conditions of the tracklets into consideration. The proposed postprocessing is particularly appropriate for tracking objects without specific appearance properties since it does not require appearance feature extraction from the objects. To prove the adequacy of the postprocessing approach, we compared the tracking performance with and without postprocessing. The postprocessing approach enhances tracking accuracy considerably. Because of the fluctuation of the captured depth information, we, finally, estimate the 3-D particle trajectories with polynomials based on the obtained 2-D trajectories to compensate for deviations.

As schematically illustrated in Fig. 1, this article is organized as follows. Section II gives an overview of the experimental setup and the plenoptic camera system. In Section III, a new methodology for particle detection combining information of 2-D gray values and 3-D point clouds is presented. We propose algorithms for particle tracking in Section IV, including a new postprocessing approach to connect tracklets into complete trajectories. Section V outlines the performed evaluation systems for both particle detection and tracking at first and discusses the results afterward.

II. Experimental Setup

In this section, we introduce the test facility BRENDA located on the Campus North of the Karlsruhe Institute of Technology. In addition, the selected camera system is described.

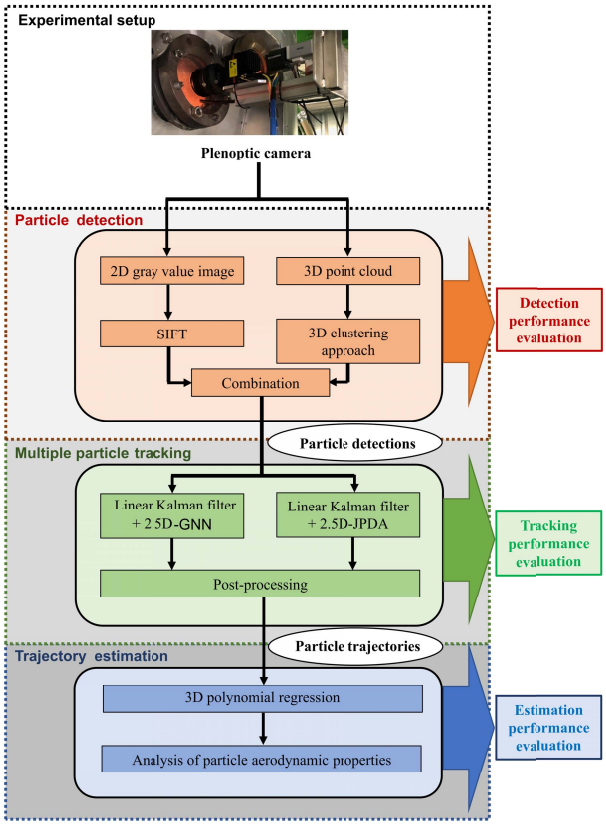


Fig. 1. Overview of 3-D RDF particle tracking-by-detection using a plenoptic camera system.

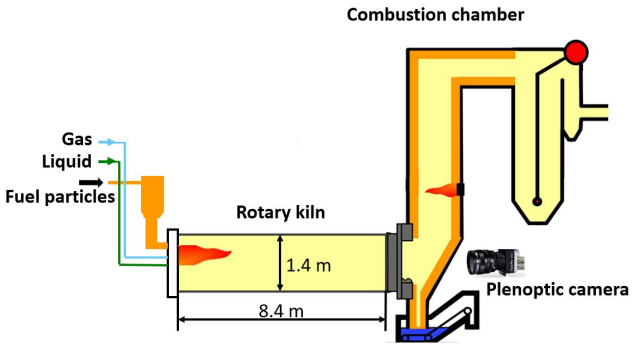


Fig. 2. Schematic of the rotary kiln and combustion chamber from the BRENDA test facility [37]. The axis of the plenoptic camera coincides with the kiln's rotation axis. The camera is fixed on a traverse in front of the glass window, as depicted in the upper part of Fig. 1. A computer at the right-hand side of the rotary kiln triggers the camera.

A. Test Facility BRENDA

The test facility BRENDA contains a rotary kiln that allows burner tests under conditions comparable to real industrial cement production. The major components of BRENDA are a rotary kiln and a postcombustion chamber, which provide a thermal power of 1.5 and 1 MW, respectively [37]. We primarily conducted experiments in the rotary kiln, which has a length of 8.4 m and an inside diameter of 1.4 m. As depicted in Fig. 2, RDF particles are conveyed through a lance into the kiln. Simultaneously, a high-speed plenoptic camera mounted

outside the kiln monitors the particles through a quartz glass window at the kiln outlet. The inside temperature of the rotary kiln can reach a maximum of 1240 °C, which causes the combustion of several fractions of RDF. In addition, the rotary kiln rotates with a velocity of 0.2 rpm, which further complicates the computing of the background images.

B. Camera System

With the development of camera technology, several cameras are capable of performing 3-D measurements, e.g., stereo camera, ToF camera, structured light camera, and plenoptic camera.

The stereo camera system is widely applied for the acquisition of stereoviews and 3-D information. Notwithstanding, a stereo camera system entails a relatively simultaneous trigger and a sophisticated calibration for both lenses. Moreover, for a stereo camera system with more than one camera, the test facility must contain more apertures and appropriate spaces for the cameras, which restricts the application of such a system.

Time-of-flight (ToF) cameras, which employ ToF techniques to acquire distance information of objects, have also found wide applications in various fields, e.g., robotics and navigations. This camera system is cost-efficient and able to provide a high real-time capability. However, the drawbacks of the ToF cameras are noticeable. Compared to other 3-D systems, the present ToF technology offers a lower resolution. Furthermore, the performance of the camera is relatively sensitive to external intrusions, such as strong ambient light or scattered light.

Structured light cameras, such as RGB-D cameras using structured light, are also a viable alternative to conduct 3-D measurements. Available RGB-D cameras, for instance, Microsoft Kinect, are significantly cost-efficient compared to other 3-D sensors. Therefore, object tracking based on RGB-D technology has been extensively researched, especially in the field of human being tracking. Nevertheless, the majority of available RGB-D cameras are not able to produce complete depth information when objects' surfaces are too thin, as small particles [38]. Moreover, the depth error of the RGB-D camera increases considerably with a distance greater than 3.5 m [39]. Hence, such cameras might not achieve sufficient accuracy for long-distance measurements.

A plenoptic camera system based on the plenoptic function is another option to realize 3-D measurements. With a microlens array in front of the image sensor, the light-field camera is able to capture the light field of the scenes, in accordance with which the 3-D information can be restored. In comparison to stereo camera systems, light-field cameras involve only a single main lens and image sensor, which benefits its universal usability. Meanwhile, the light-field cameras precede the ToF cameras in terms of resolution and robustness. Besides, the accuracy and adequacy of the focused plenoptic camera for long-distance measurements are proved in [40].

Considering comprehensive factors, e.g., the construction of the test facility (merely single optic access feasible and the distance between particles and the camera could reach above 8 m), and sufficient resolution to capture small particles with enough visibility, we selected a plenoptic camera system to capture the fuel particles.

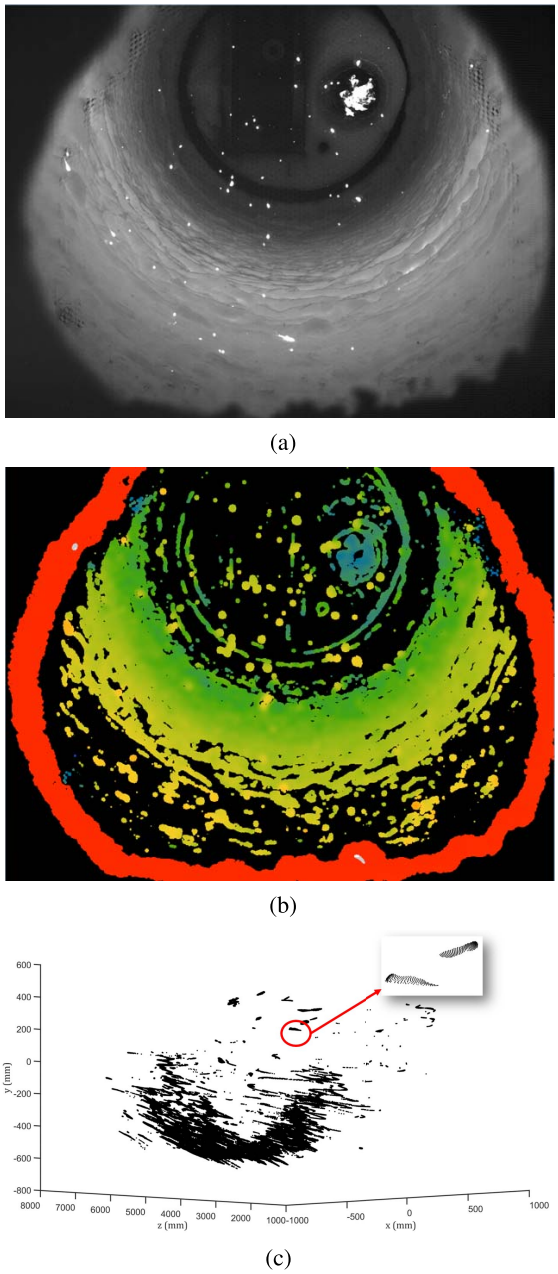


Fig. 3. Example of the captured image. (a) Total focus image. The captured image corresponds to the image captured by a conventional camera. (b) Depth map. The depth information is depicted as a false-color image. Black indicates no depth information available. (c) Point cloud. The camera is able to convert the captured image information into point clouds.

The utilized plenoptic camera from Raytrix provides a resolution of 1536 pixel × 2048 pixel at a framerate of 330 frames/s (fps). The lens used has a focal length of 85 mm. Besides, the microlens array of the applied camera consists of microlenses with different focal lengths, which achieves a wide range of depth of field and a high maximal lateral resolution [41]. Fig. 3(a) shows an example of the captured gray-scale image, and Fig. 3(b) shows the corresponding 3-D-depth map.

As shown in Fig. 3(b), not every pixel can be successfully captured spatially with specific depth information. Thus, the total amount of points is much less than the resolution.

As illustrated in Fig. 3(c), point clouds are also available, where each captured pixel of a 2-D image owns a spatial position with x-, y-, and z-coordinates.

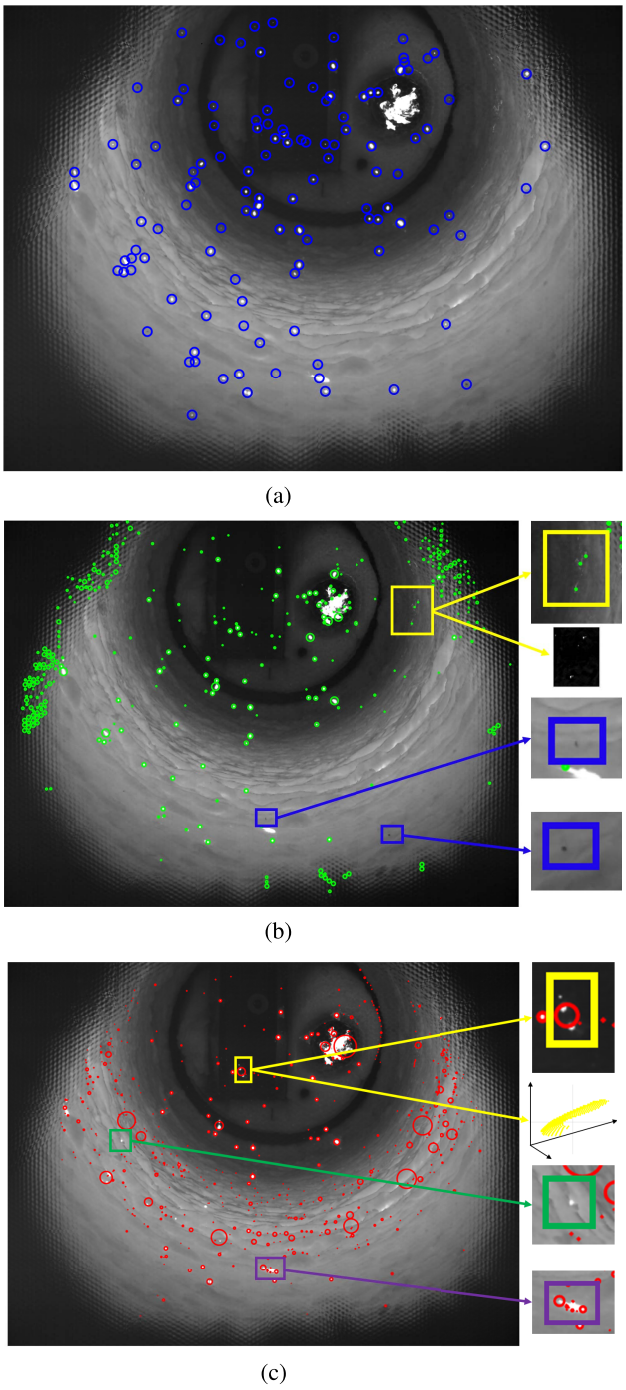
III. Particle Detection

In this section, we elaborate on a novel particle detection approach, which combines the detection results of two distinct algorithms based on different concepts to facilitate detection performance.

As described in Section II, the applied plenoptic camera stores both the gray value and the available spatial position of the pixels. Based on the gray value images, several 2-D detection approaches can identify the fuel particles, as mentioned in Section I. Here, we implemented a preprocessing of background subtraction for the 2-D-SIFT algorithm [5] to reduce the impact of the background. The background model is received by computing the temporal median gray value for each pixel within a time period (generally, 75 previous frames and 75 posterior frames). Because of the slow rotation of the rotary kiln, the selected filter time period for the background estimation ought to be chosen long enough to exclude fast-moving RDF particles and also short enough to avoid distortion of the background model. Meanwhile, the particles can also be detected by forming clusters in the obtained point clouds using clustering methods. In this article, we applied three clustering algorithms: QCC, MO-based clustering, and DBSCAN. Here, we use the DBSCAN approach as an example of the 3-D clustering algorithms.

We primarily applied both detection concepts to the collected datasets. To compare the results visually, we convert the clustering results into 2-D by computing the clusters' gray value weighted central pixel position. In addition, we manually detect the particles (ground truth) and mark the positions for further quantitative detection performance evaluation. Fig. 4 shows the detection results of 2-D-SIFT and 3-D-DBSCAN together with the labeled ground truth.

Apparently, substantial numbers of false detections occur in the edge regions of the image owing to artifacts caused by the camera. This issue will be addressed later in this article by introducing an appropriate region of interest (ROI) that includes as many particles as possible and, meanwhile, excludes the artifacts region. Despite the defined ROI, both detection results in Fig. 4 show particular demerits. The gray value-based 2-D-SIFT algorithm is not able to detect particles with slight brightness, as marked in the blue boxes in Fig. 4(b). In addition, due to disturbances in the estimated background image, gray value peaks may occur in the foreground image after background subtraction, where no particle is present, as marked in the yellow box in Fig. 4(b). Since background subtraction is a universal preprocessing for object detection to reduce interferences from the background, the therefrom occurred detection errors constitute a substantial obstacle for detection. Unlike gray value-based detection approaches, clustering methods are less sensitive to these luminance issues, whereas the performances of the clustering methods are constrained by other factors, for instance, the distribution of the points within the cloud. For one relatively big particle, several clusters might be formed inside the particle due to the unfavorable point distribution, as marked in the purple box in Fig. 4(c). In addition, several particles could be identified as one cluster since their corresponding point clouds are



(a)

(b)

(c)

Fig. 4. Result of particle detection. (a) Manually labeled ground truth. Each blue circle indicates one particle. (b) Detection result of 2-D-SIFT. The center of each green circle corresponds to the position of the detected particle, whose sizes are pointed by the circles' radius. The yellow box is zoomed in to display the detection on the initial frame and the frame after background subtraction. (c) Detection result of 3-D-DBSCAN. The center of each red circle corresponds to the center of the cluster. The radius of the circle indicates the cluster size. The yellow box shows the particle detection both on the 2-D frame and in the 3-D point cloud.

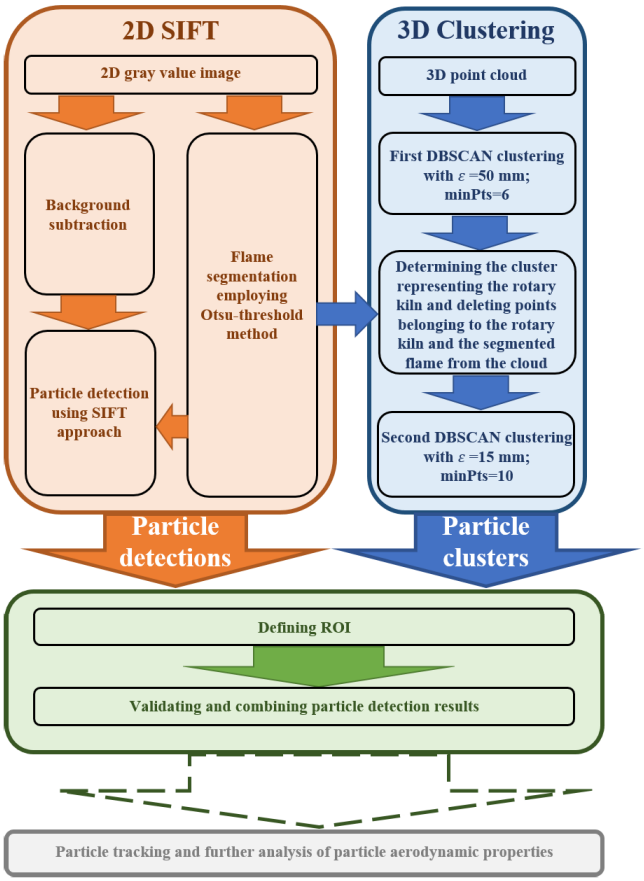


Fig. 5. Schematic of the presented combined particle detection approach.

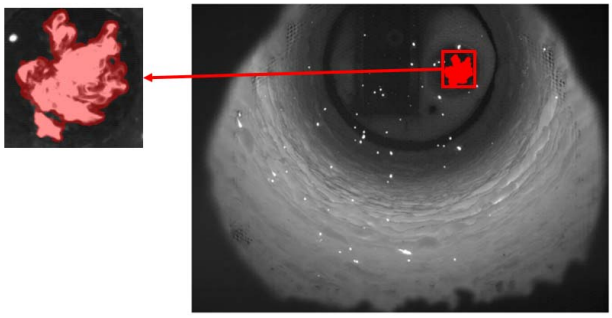


Fig. 6. Flame segmentation within a defined rectangular ROI around the burner, as marked in red rectangular. The red-marked flame on the left is the segmented flame.

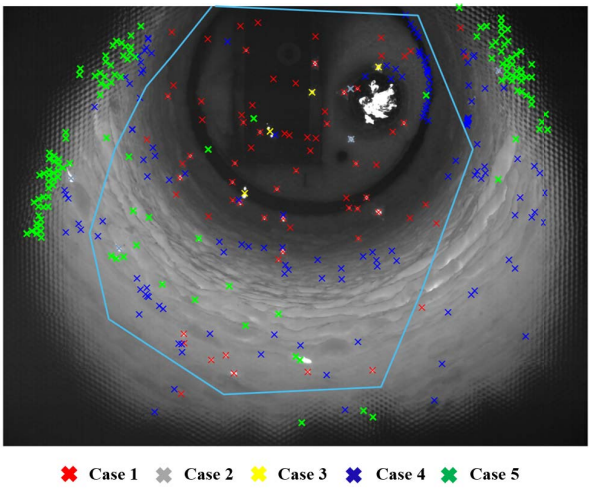


Fig. 7. Five cases of the comparison of particle detections. The blue polygon is an instance of the defined ROI.

connected spatially, as shown in the yellow box in Fig. 4(c). Similarly, particles lying on the wall can be incorporated into the cluster of the rotary kiln leading to missing detections, as in the example marked in the green box.

Under this circumstance, we propose a novel combined detection approach, which synthesizes the gray value-based algorithm 2-D-SIFT and the 3-D-clustering method DBSCAN. The presented approach is schematically illustrated in Fig. 5. We separately implement 2-D-SIFT and 3-D-clustering first. As mentioned, for preprocessing, the background can be modeled by computing the temporal median gray value for each pixel within a specific time period. Alternatively, we could also use the difference of successive frames to acquire simple foreground models. To avoid negative influences of the flame on particle detection, we also segment the flame inside the image before detection. The flame segmentation within a certain image is accomplished by utilizing the OTSU threshold selection approach [4] on a defined rectangular ROI around the burner, as illustrated in Fig. 6. Since the gray value of the flame differs considerably from the background, the flame region can be segmented efficiently and effectively by the OTSU approach.

As indicated in Fig. 5, the clustering method DBSCAN is employed twice with distinct parameter values. In accordance with the definition of DBSCAN [19], ε stands for the size of the search radius around a point, and $minPts$ represents the threshold of the point amount within the search radius for a point to be deemed a core point. A large ε and a small $minPts$ benefit the formation of big clusters, e.g., the rotary kilns inner wall, while, on the contrary, a small ε and a sizeable $minPts$ help distinguish particles accurately. After implementing the first clustering, the cluster representing the rotary kiln is determined, and the corresponding points together with the points of the segmented flame are removed from the point cloud to facilitate the performance of the second clustering. With the completion of these steps, we are now ready to combine the detection results.

First, an ROI is defined to exclude the peripheral artifacts in the image. Subsequently, the detection results of SIFT and DBSCAN are compared and validated with each other. For ease of comparison, the formed 3-D clusters are converted into 2-D by presenting the points in pixel coordinates. Taking the particle detections with respect to their sizes into consideration, we can conclude the following five cases for the comparison, as schematically depicted in Fig. 7.

1) *Case 1:* A particle candidate is identified as exactly one particle by both methods.

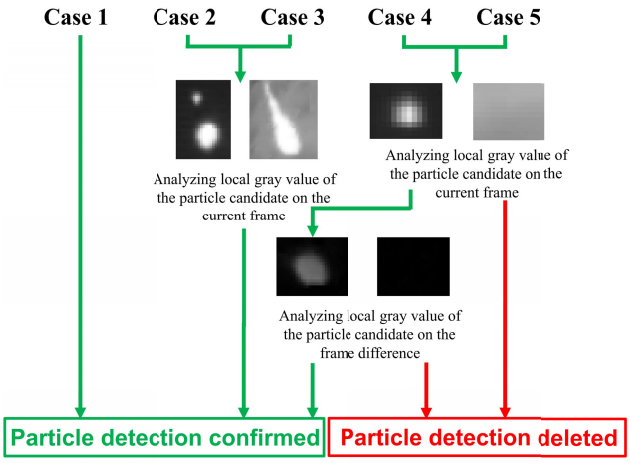


Fig. 8. Approach to dealing with various detection cases.

2) *Case 2:* A particle candidate is detected as one particle by the SIFT algorithm but is identified as more than one particle by the clustering approach.
3) *Case 3:* A particle candidate is detected as one particle by the clustering approach but is identified as more than one particle by the SIFT algorithm.
4) *Case 4:* A particle candidate is only discovered by the clustering approach.
5) *Case 5:* A particle candidate is only detected by the SIFT algorithm.

For the sake of dealing with the mentioned cases, we recommend the logic demonstrated in Fig. 8. If both methods validate one particle detection, this detection will be directly accepted in the result (case 1). In case of different detection outcomes, as in case 2 and case 3, we will analyze the gray value distributed on the entire particle area of the current frame to determine the precise amount of particles. On the condition that the particle detection is merely confirmed by one approach (cases 4 and 5), the gray value of the particle area on both the current frame and the frame difference computed by subtracting the previous frame from the current frame will be considered. Only if the gray value distribution indicates a particle-like feature in both situations, the particle will be finally added to the detection list.

IV. Particle Tracking and Postprocessing

In this section, we detail the complete 2-D tracking procedure, including a postprocessing approach on the basis of the particle motion similarity that we first introduced in [42].

A. 2.5-D Multiple-Particle Tracking

In order to associate the particle detections and obtain preliminary 2-D particle tracklets, we perform two widely used tracking algorithms and the Kalman filter with GNN and JPDA two-dimensionally in pixel coordinates. Both algorithms follow a similar functioning pattern: prediction, gating, assignment, and update. Because of the high-speed plenoptic camera with up to 330 fps, the particle movements within a small number of consecutive frames can be approximated as a+ uniform movement. Thus, we apply a linear Kalman Filter containing a uniform movement model for the prediction of the particle positions. In order to restrict potential detection candidates for a certain prediction, we compute the 2-D Mahalanobis ellipse according to a predefined distance threshold around the prediction as a 2-D gating region. Moreover, we also regard depth changes as a gating factor. Only if detection is within the ellipse of a unique prediction and the change in depth compared to the corresponding detection in the previous image does not exceed a defined limit, the detection will be judged as a possible candidate.

For each candidate, the assignment cost will be computed utilizing the function presented in [28]. Since we take the depth information into account, we, therefore, modify the tracking approach to 2.5-D. In GNN, the whole computed assignment costs constitute the cost matrix, according to which a biunique association of detections and tracks is performed using the Kuhn–Munkres algorithm [43]. Unlike GNN, JPDA considers all detections within the gate and weights them for further steps instead of a biunique association and discarding the others.

Finally, the tracks are updated in terms of the association outcomes, as demonstrated in [28] and [30]. Our approach considers all thereby obtained particle traces as tracklets regardless of their actual lengths and durations.

Principally, we could also integrate the depth changes into the cost function to enhance assignment accuracy. Notwithstanding, the depth information provided by the camera fluctuates significantly and is, thus, not reliable and appropriate to be added to the cost function quantitatively.

B. Postprocessing

After the particle detection and tracking have been carried out successively, the analysis of particle aerodynamic properties is able to proceed in principle. However, the straightforwardly acquired tracklets fail to provide sufficient reliability due to the presence of tracking errors, such as incomplete tracklets and false tracklets, which occur primarily on account of detection inaccuracy. Hence, we propose and demonstrate a postprocessing framework for multiple-particle tracking aiming at enhancing tracking performance, as schematically illustrated in the flowchart in Fig. 9. In the following, the presented steps of the framework are described gradually.

1) Faulty Tracklets Elimination: As the first step, tracklets that are extremely short, moving quite slowly, or too irregularly are regarded as faulty tracklets and, therefore, ought to be eliminated from the tracklet list. Only tracklets that fulfill the condition

$$\mathcal{D}(T_i) > \mathcal{D}_\epsilon \quad \& \quad |\overline{V}_i| > |\overline{V}|_\epsilon \quad \& \quad \mathcal{R}_{T_i\widetilde{T}_i} > \mathcal{R}_\epsilon \tag{1}$$

are in the focus of interest and remain for further steps. Here, $\mathcal{D}(T_i)$ denotes the duration of tracklet T_i, $|\overline{V}_i|$ stands for the arithmetic average magnitude of the temporary velocity vector of T_i, and $\mathcal{R}_{T_i\widetilde{T}_i}$ computes the cross correlation coefficient of T_i and $\widetilde{T}_i$, which approximates T_i with a second-degree polynomial. In addition, the subscript ϵ indicates the respective threshold value of the corresponding parameter.

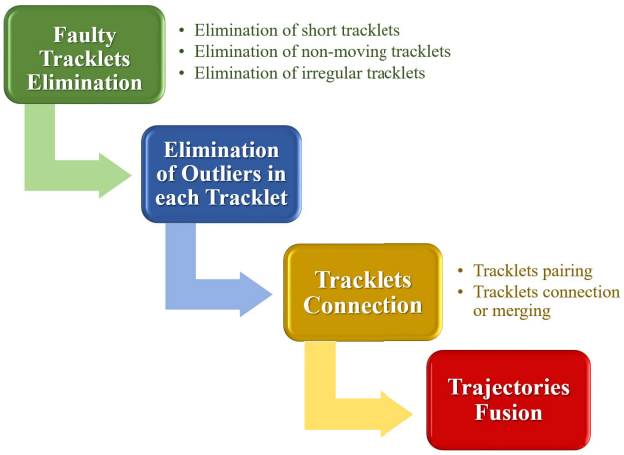


Fig. 9. Outline of the postprocessing framework.

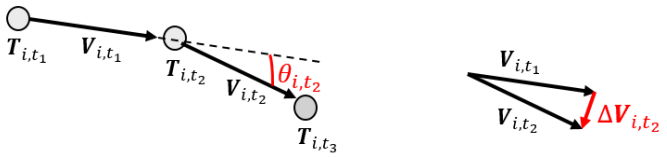


Fig. 10. Schematics of the notations for outlier detection.

2) Elimination of Outliers in Each Tracklet: Taking the fact into account that false detections could persist in retained tracklets, we decide to identify and delete outliers in the tracklets. The majority of the outliers are a consequence of detected noise and result in local irregular peaks in the tracklets. Thus, we consider a detection belonging to T_i at time t as an outlier if

$$\theta_{i,t} > \theta_\epsilon \quad \& \quad |\Delta V_{i,t}| > |\Delta V|_\epsilon \tag{2}$$

where $\theta_{i,t}$ is the angle between two successive velocity vectors and $|\Delta V_{i,t}|$ denotes the magnitude of the difference between them, as shown in Fig. 10.

3) Tracklet Connection: After the previous two steps accomplish the preparation, we are able to process the vital task of tracklet connection. As depicted in Fig. 11, tracklet connection is an iterative process, which starts with scanning for possible tracklet matching pairs among the entire tracklet list. The corresponding pairing costs will be computed afterward and construct the cost matrix accordingly. Based on the cost matrix, the connection of bilateral NN matching pairs is executed, and the existing tracklet list will be updated subsequently. The process ends until there is no possible tracklet matching pair.

Before demonstrating the components in Fig. 11 separately, we pinpoint two cases in tracklet connection in terms of time overlap with Fig. 12. If the potential matching pair do not overlap in the timeline, we define their connection as a simple connection. On the contrary, if the matching pair appear partially simultaneously, their connection is associated with merging.

In Fig. 12, we also define the concept of the connected part T_{ik}, which is underlined by black dashed boxes. For a simple

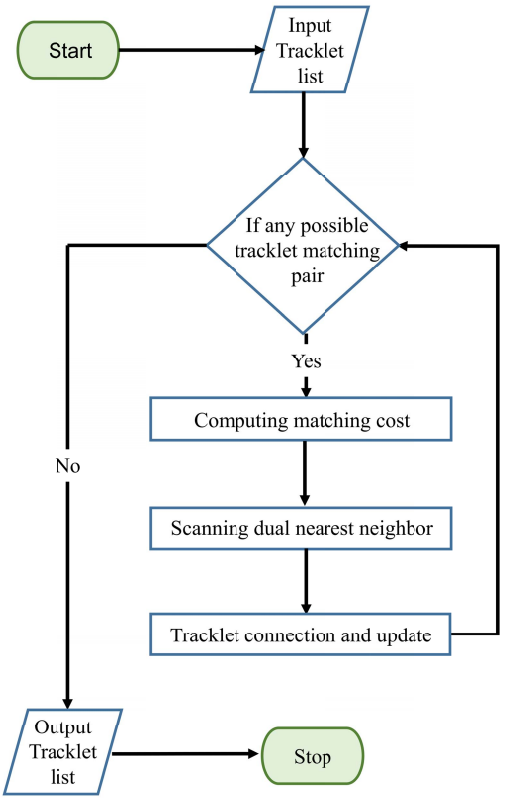


Fig. 11. Flowchart of tracklet connection.

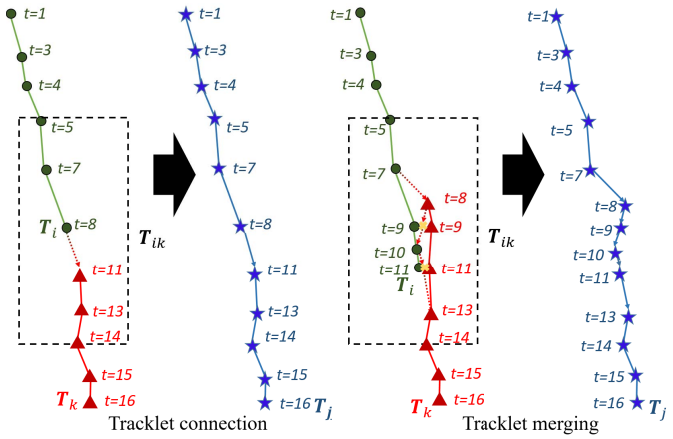


Fig. 12. Tracklet connection and tracklet merging. T_i and T_k are possible matching pair that is connected or merged to build a new tracklet T_j. Left: connection of tracklets T_i and T_k. Right: tracklet merging. Dashed rectangular marks the connected part T_{ik}. For tracklet connection, the tracklets are simply connected, whereas, in tracklet merging, the midpoints of simultaneous detections in both tracklets are first computed and substitute the original positions in the built tracklet T_j.

tracklet connection, the connected part consists of three last temporary detections of the antecedent tracklet together with three first detections of the subsequent tracklet. For tracklet merging, the first temporally coincident detection and its two previous detections of the antecedent tracklet, as well as the last temporally coincident detection and its following two detections from the subsequent tracklet, belong to the connected part.

A time gap exists in tracklet connection, while time overlap is a prerequisite for tracklet merging. In order to be a possible tracklet matching pair, the time gap or time overlap of the tracklet candidates ought to vary in a rational range, namely,

$$t_{T_k,1} - t_{T_i,\text{end}} < t_{\epsilon,\text{connect}} \quad \text{or} \quad n_{T_i} \cap n_{T_k} < n_{\epsilon,\text{merge}}. \tag{3}$$

Thereby, we follow the notations in Fig. 12, where T_i denotes the antecedent tracklet and T_k stands for the subsequent tracklet. $t_{T_k,1}$ is the time when T_k begins, while $t_{T_i,\text{end}}$ denotes the end time of T_i. In addition, n is a set of discrete appearance times of a particular tracklet, and t_ϵ and n_ϵ are predefined thresholds with subscripts, indicating different values for tracklet connection and merging, respectively. $t_{\epsilon,\text{connect}}$ constrains the time gap of the pairing candidate, and $n_{\epsilon,\text{merge}}$ indicates that the paring tracklets are not allowed to have the simultaneous time cardinality overlap greater than a threshold.

Besides the time restriction in (3), the matching pair should also satisfy the following two requirements: first, no outliers in the connected part; second, the connected part itself is not a faulty tracklet according to (1) but without the limitation of duration. After determining potential tracklet matching pairs, their pairing costs are calculated by

$$\begin{aligned}\mathcal{C}(T_i, T_k) = a\overline{\theta}_{ik} + b|\overline{\Delta \boldsymbol{V}}_{ik}| + c\overline{\Delta z}_{ik} + e|n_{T_i} \cap n_{T_k}| \\ + d \cdot \max\left(\left(t_{T_k,1} - t_{T_i,\text{end}}\right), 0\right) + \frac{f}{\max\left(\overline{F}_{ik}, 1\right)}\end{aligned} \tag{4}$$

where $\overline{\theta}_{ik}$ denotes the arithmetic average of $\theta_{ik,t}$ with the unit degree, $|\overline{\Delta \boldsymbol{V}}_{ik}|$ is the arithmetic average of $|\Delta \boldsymbol{V}_{ik,t}|$ with the unit pixel per frame, and $\overline{\Delta z}_{ik}$ stands for the arithmetic depth change between frames with the unit millimeter per frame. These first three terms are conducive to smooth connections. The fourth and fifth terms penalize the matching pairs with long time gaps or time overlaps. Thereby, $|n_{T_i} \cap n_{T_k}|$ denotes the number of elements of the intersection of the sets of appearance time. $\overline{F}_{ik}$ in the last term is derived from the laws' properties, also called "texture energy measures" [44], which pursues the goal of texture feature extraction by filtering images with five types of masks: level, edge, spot, ripple, and wave. By convolving a texture image with combinations of these masks, the microstructure of the texture will be accentuated [45]. Comparing diverse combinations of these masks, we select the level-edge mask that delivers considerable difference in values with and without tracklet-like features in the images. We apply the laws' mask to a window on the sum of frame differences obtained by adding a set of consecutive frame difference images within the time period of the tracklets' appearance. The window is defined as a parallelogram with the following vertexes for a vertically dominant connection:

$$\begin{aligned}&\left(T^c_{i,t_{T_i,\text{end}}} - w, T^r_{i,t_{T_i,\text{end}}}\right), \left(T^c_{i,t_{T_i,\text{end}}} + w, T^r_{i,t_{T_i,\text{end}}}\right) \\ &\times \left(T^c_{k,t_{T_k,1}} - w, T^r_{k,t_{T_k,1}}\right), \left(T^c_{k,t_{T_k,1}} + w, T^r_{k,t_{T_k,1}}\right).\end{aligned}$$

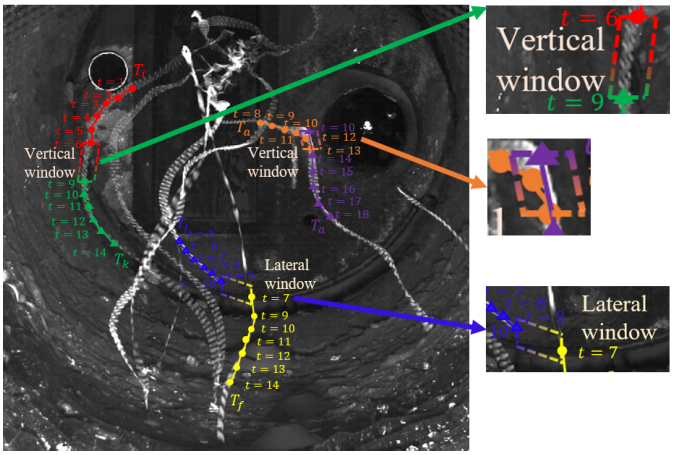


Fig. 13. Example of windows for laws' mask on the sum of difference images acquired by background subtraction. The left window (red/green) is a vertical window for tracklet connection, where the connection is primarily vertical. Correspondingly, the right window (blue/yellow) is a lateral window for tracklet merging. Moreover, the window (orange/purple) indicates a vertical window for tracklet merging.

Accordingly, a lateral connection entails a parallelogram with the vertexes

$$\begin{aligned}&\left(T^c_{j,t_{T_j,\text{end}}}, T^r_{j,t_{T_j,\text{end}}} - w\right), \left(T^c_{j,t_{T_j,\text{end}}}, T^r_{j,t_{T_j,\text{end}}} + w\right) \\ &\times \left(T^c_{f,t_{T_f,1}}, T^r_{f,t_{T_f,1}} - w\right), \left(T^c_{f,t_{T_f,1}}, T^r_{f,t_{T_f,1}} + w\right).\end{aligned}$$

Thereby, $T^c_{i,t}$ denotes the column position of the tracklet T_i at time t, and $T^r_{i,t}$ stands for the row position of the tracklet at time t, for instance, $T^c_{i,t_{T_i,\text{end}}}$ is the column position of tracklet T_i at the end time of the tracklet T_i. w is half of the window's length. The direction of the window is selected regarding the dominant connection distance of the pairing tracklets. Fig. 13 shows examples of windows and the sum of different images. It is obvious that the window size of a particular matching case depends on the locations of the corresponding tracklets and the window length w, which is selected according to the average particle size in the image coordinate. Using laws' properties instead of a simple sum of gray values [42] reduces the impacts of accidental illumination changes and enhances the robustness afterward.

In (4), a, b, c, d, e, and f denote the corresponding weighting factors of each term, whose values are discussed in Section V. To process the connection in accordance with the computed pairing cost from (4), we recommend the dual NN (DNN) algorithm. One possible tracklet matching pair is allowed to be connected only if they are the bilateral NN, namely, the one tracklet is the NN of the other, and meanwhile, the other one is also the NN of the first one. From the performance aspect, the DNN algorithm precedes the Kuhn–Munkres (GNN) algorithm, which seeks a sole optimal solution that might lead to an inaccurate connection.

We highlight the advantage of DNN using the example in Fig. 14. In the cost matrix, the connection solution in each iteration is underlined by orange background for the Kuhn–Munkres algorithm and marked in the green box for DNN. As depicted in Fig. 14, the Kuhn–Munkres algorithm

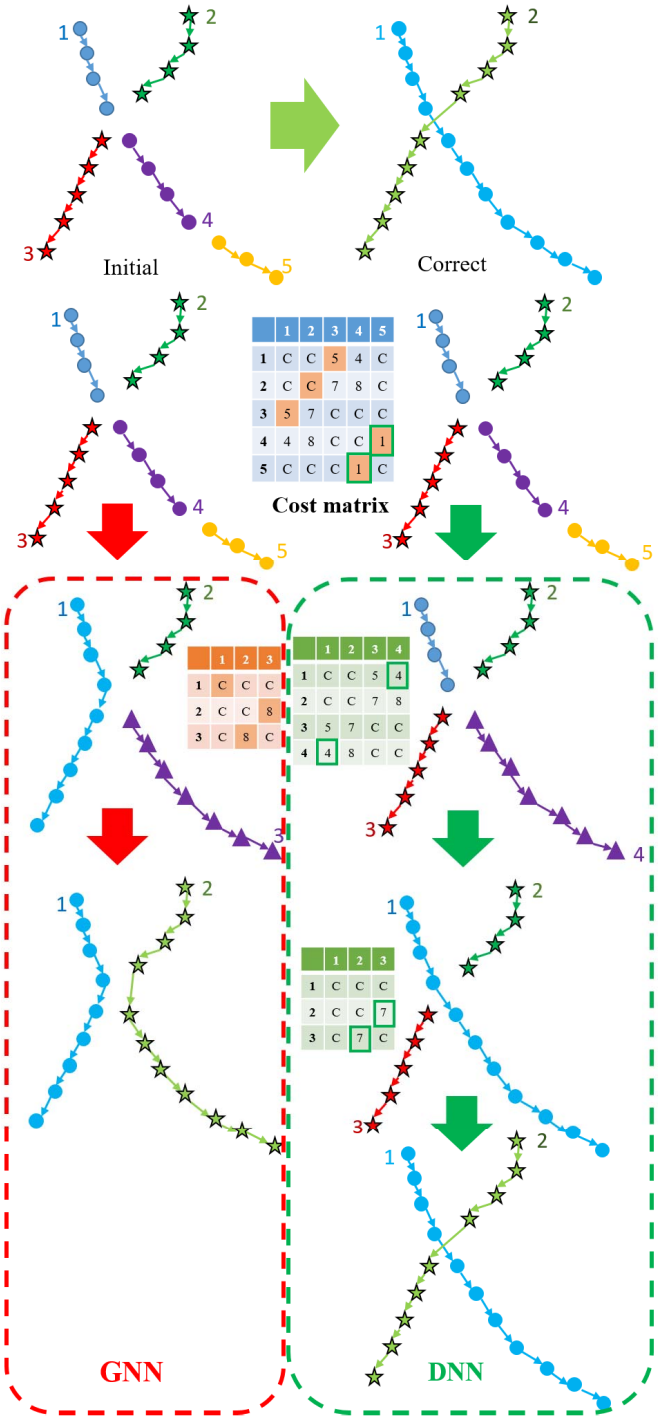


Fig. 14. Example of tracklet connection using the Kuhn–Munkres algorithm and DNN, respectively. In the cost matrix, C indicates the impossibility of a connection pair with an extraordinarily large value.

wrongly connects tracklets 1 and 3 in order to reach the global minimal cost and results in an undesired outcome as a consequence. In comparison to the Kuhn–Munkres algorithm, DNN connects merely the DNN matching pair in each iteration and succeeds in obtaining the expected result.

4) Tracklet Fusion: With the completion of the last three steps, the rest of the rational tracklets are connected into trajectories. For the majority of applications, the task is accomplished. Nevertheless, the abovementioned three steps might not be sufficient to achieve satisfactory accuracy when dealing with detections containing many duplications and noise since, under this circumstance, one particle might trigger more than one trajectory within a certain period. Thus, we suggest a further step trajectory fusion that differs slightly from tracklet merging. Principally, tracklet merging and trajectory fusion follow the identical procedure. The only difference in trajectory fusion is that the requirements in (3) and terms concerning the time gap and overlap in (4) are omitted.

TABLE I

PERFORMANCE EVALUATION OF PARTICLE DETECTION VIA CLUSTERING METHODS, 2-D-SIFT WITH MEDIAN BACKGROUND SUBTRACTION, AND THE NOVEL COMBINED APPROACH

	TP	FP	FN	Precion	Recall	F_1-score
QCC Clustering	3260	5242	1127	0.3834	0.7431	0.5058
MO-based Clustering	2869	4389	1518	0.3953	0.6540	0.4928
DBSCAN Clustering	3175	4445	1212	0.4167	0.7237	0.5289
SIFT	3687	354	700	0.9124	0.8404	0.8749
Combi SIFT+DBSCAN	3797	229	590	0.9431	0.8655	0.9026

V. EVALUATION SYSTEMS AND RESULTS

In this section, we summarize the results of RDF particle detection and tracking using the presented approaches. The applied quantitative evaluation systems are briefly illustrated as well.

A. Particle Detection

To proceed with an objective assessment of the detection performance, we manually identify and label particles of fifty frames from five different periods with distinct properties as ground truth. Since the ground truth is only available in 2-D pixel coordinates, the evaluation is processed two-dimensionally as well by representing each particle with its central pixel position. The detections are assigned to simultaneous particles in the ground truth by the Kuhn–Munkres algorithm. Considering the particle size, we select the value of 15 pixels as the maximal cost in the corresponding cost matrix. The successfully assigned detections are denoted as true positive (TP), and on the opposite, the detections that are failed to associate with particles in the ground truth are considered false positive (FP). False negative (FN) refers to the nonassigned particles in the ground truth. The precision, recall, and F1-score are computed accordingly. The ground truth is divided into five datasets in accordance with their time dimension, and we apply the cross-validation technique to the datasets for determining the threshold value in SIFT. The results are illustrated in Table I.

It is obvious that the 3-D-clustering approaches provide the lowest precision, whose value is barely half of the SIFT and combi methods, since the results contain more nonparticle detections, such as the rotary kiln. Moreover, the recalls of the clustering approaches are bottom as well, which results in unsatisfactory F1-scores. Generally, the three applied clustering methods deliver comparable outcomes. MO-based clustering fails to detect enough TPs, and thus, its recall ranks at the

bottom. QCC is able to identify more TPs. Nevertheless, QCC detects more nonparticle objects as well, which leads to low precision. DBSCAN could discover sufficient TPs and avoid FPs simultaneously, which brings the highest F1-score among the clustering methods. The average F1-score of the clustering approaches is around 0.5. The unsatisfactory outcomes are the consequence of the unsteady 3-D point clouds provided by the camera that captures both particles and clutters. Since the DBSCAN ranks at the top with respect to the F1-score, the following combi result and detection result for further tracking are based on DBSCAN.

Compared to the clustering approaches, the 2-D-SIFT algorithm reaches a high precision value of 0.9124, which indicates a detection result with less than 10% false detections (noise and duplications). The novel combined method owns the highest recall of 0.8655, representing a misdetection rate of around 13.5%. These misdetections are generally small particles with low brightness and without spatial positions, which both 2-D-SIFT and 3-D-clustering cannot detect. Despite the combination with the 3-D-clustering approach that introduces substantial numbers of false detections, the precision of the combined method is increased to 0.9431. Consequently, the combined method achieves the best F1-score with 0.9026. In terms of the amount of TP, FP, and FN, the combination approach detects 110 more TPs and 125 fewer FPs. In addition, roughly, 450 particles in the ground truth are small gloomy particles without 3-D information that limits the upper recall value of all the utilized approaches.

B. Particle Tracking

Analogously, the tracking performance is quantitatively evaluated by a tool recommended in [46]. The evaluation tool matches the obtained trajectories with the ground truth that is also manually labeled for the work and computes several defined measurements accordingly. Here, four measurements are vital for performance assessment, namely, normalized pairing score α, full normalized score β, the Jaccard similarity index for positions JSC, and the Jaccard similarity index for tracks JSC_θ. According to the definition in [46], the measurements of a set of candidate tracks $\mathcal{Y}$ and a set of ground-truth tracks $\mathcal{X}$ are specified as follows.

1) *Normalized Pairing Score α:*

$$\alpha(\mathcal{X}, \mathcal{Y}) \triangleq 1 - \frac{d(\mathcal{X}, \mathcal{Y})}{d(\mathcal{X}, \emptyset)}. \tag{5}$$

Hereby, $\emptyset$ stands for the set of $|\mathcal{X}|$ dummy empty trajectories and $d((\mathcal{X}, \emptyset)$ is the distance between $\mathcal{X}$ and $\mathcal{Y}$.

2) *Full Normalized Pairing Score:*

$$\beta(\mathcal{X}, \mathcal{Y}) \triangleq \frac{d(\mathcal{X}, \emptyset) - d(\mathcal{X}, \mathcal{Y})}{d(\mathcal{X}, \emptyset) + d\left(\bar{\mathcal{Y}}, \emptyset_{\bar{\mathcal{Y}}}\right)} \tag{6}$$

where $\bar{\mathcal{Y}}$ denotes the trajectories in $\mathcal{Y}$ but not in $\mathcal{Z}^*$ and $\emptyset_{\bar{\mathcal{Y}}}$ stands for a set of dummy trackjectories with $|\bar{\mathcal{Y}}|$ elements. $\mathcal{Z}^*$ is the obtained ensemble set of tracks by choosing $|\mathcal{X}|$ tracks from the extended set of candidate tracks that yield the minimal distance to $\mathcal{X}$.

Fig. 15. Various investigated fuel fractions. (a) Wood chips. (b) Confetti. (c) Paper shreds. (d) PE granules.

3) *Jaccard Similarity Index for Positions:*

$$\text{JSC} \triangleq \frac{\text{TP}}{\text{TP} + \text{FN} + \text{FP}}. \tag{7}$$

Tp denotes the number of matching pairs of positions in $(\mathcal{X}, \mathcal{Z}^*)$. FN stands for the number of nonmatching pairs. FP represents the total amount of positions in $\bar{\mathcal{Y}}$.

4) *Jaccard Similarity Index for Tracks:*

$$\text{JSC}_\theta \triangleq \frac{\text{TP}_\theta}{\text{TP}_\theta + \text{FN}_\theta + \text{FP}_\theta}. \tag{8}$$

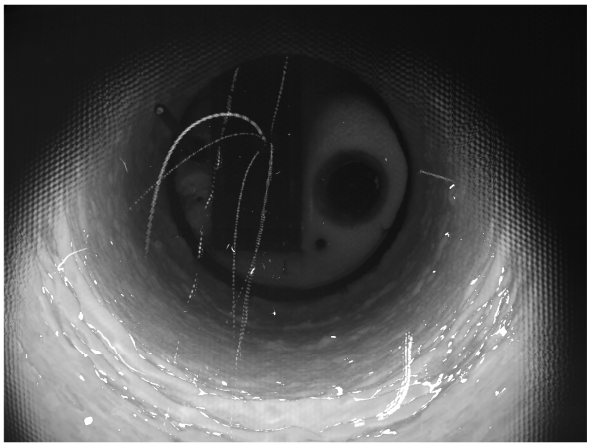
TP_θ denotes the number of nondummy tracks in $\mathcal{Z}^*$. On the contrary, FN_θ denotes the number of dummy tracks in $\mathcal{Z}^*$. FP_θ counts the number of tracks in $\bar{\mathcal{Y}}$.

The higher the scores, the better the tracking performance is.

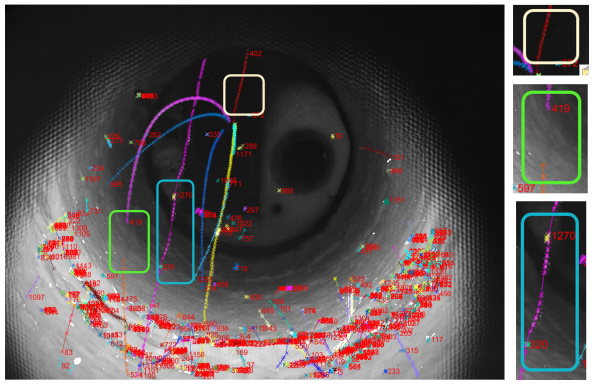
Unlike the results of particle detection, where the RDF mixtures are conveyed into the test facility for the purpose of creating a comprehensive detection scenario, for particle tracking, various fuel particles are conveyed into the rotary kiln separately to investigate the respective flight behaviors of each fuel. Overall, four distinct fuels are investigated: wood chips, confetti, paper shreds, and PE granules, as shown in Fig. 15.

First, we process 2.5-D particle tracking with GNN and JPDA. To realize a preliminarily qualitative and visual evaluation of the tracking results, we add a set of difference images to a single image as optical ground truth, as depicted in Fig. 16(a).

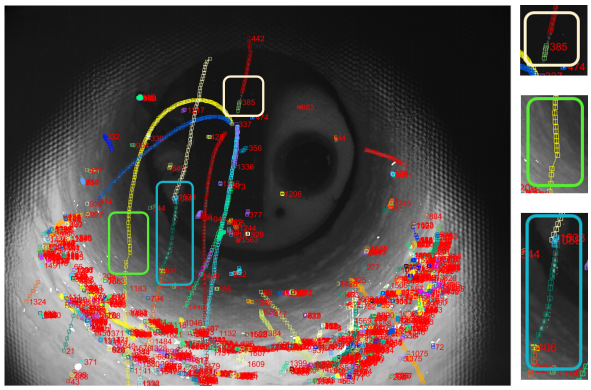
The tracking result of GNN and JPDA without postprocessing is shown in Fig. 16(b) and (c), respectively. The identical color indicates an identical tracklet. Since the computational



(a)



(b)



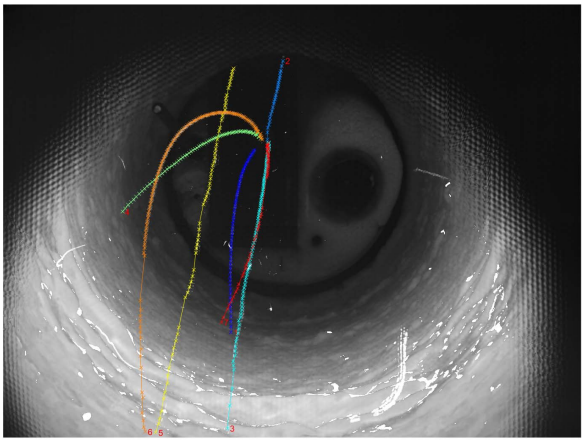
(c)

Fig. 16. Optical ground truth and tracking results of wood chips. (a) Optical ground truth. (b) Tracking result of GNN. Crosses stand for detections. Detections of the same tracklets are linked with solid lines. (c) Tracking result of JPDA. Squares stand for detections. Detections of the same tracklets are linked with solid lines.

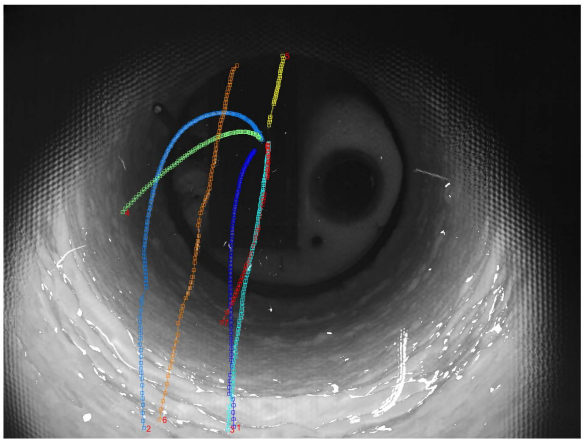
cost and entailed storage capacity of JPDA increase enormously with the rising gate and the number of tracks, the gate of JPDA is selected to be smaller than GNN. While the updated variance by JPDA is much larger than GNN owing to an additional term according to the mixture probability density function [29], [30], the gating ellipse determined by JPDA is not necessarily smaller than GNN. It is obvious that discrepancies exist in the tracking results, as highlighted by the boxes in Fig. 16. These discrepancies are primarily caused by different values of threshold and distinct functions for updating the covariance matrices.

TABLE II
PARAMETER VALUE

Parameter	Value	Unit	Parameter	Value	Unit
$\mathcal{D}_\epsilon$	2	frame	$t_{\epsilon,1}$	15	frame
$\lvert\boldsymbol{V}\rvert_\epsilon$	1	pixel/frame	$t_{\epsilon,2}$	5	frame
$\mathcal{R}_\epsilon$	0.9	-	z_ϵ	1500	mm
$\lvert\Delta\boldsymbol{V}\rvert_\epsilon$	18	pixel/frame	ϵ	10	pixel
θ_ϵ	30	$^\circ$	a	2	-
b	2	-	c	0.1	-
d	1	-	e	2	-
f	10^{11}	-	w	15	pixel



(a)



(b)

Fig. 17. Tracking results of wood chips with postprocessing. (a) Tracking result of GNN. Crosses stand for detections. Detections of the same tracklets are linked with solid lines. (b) Tracking result of JPDA. Squares stand for detections. Detections of the same tracklets are linked with solid lines.

In general, the direct outcomes of both approaches fail to provide sufficient accuracy for further analysis of spatial particles' flight trajectories and properties. Hence, we apply the presented postprocessing framework to the tracking results with the parameter values listed in Table II.

The tracking results with postprocessing using GNN and JPDA differ only slightly from each other, as schematically illustrated in Fig. 17. Compared to tracking results without postprocessing in Fig. 16, the tracking outcomes with postprocessing show better visual accuracy. For the purpose of a quantitative evaluation and comparison of the tracking performance, we assessed the tracking results utilizing the tool

TABLE III

EVALUATION OF TRACKING PERFORMANCE WITH AND WITHOUT POSTPROCESSING FROM GNN AND JPDA. UPPER VALUES ILLUSTRATE THE PERFORMANCE OF GNN AND JPDA WITHOUT POSTPROCESSING, AND LOWER VALUES DISPLAY THE PERFORMANCE OF GNN AND JPDA WITH POSTPROCESSING

Fraction \ Measure	α		β		JSC		JSC_θ	
	GNN	JPDA	GNN	JPDA	GNN	JPDA	GNN	JPDA
Wood chips	0.61	0.61	0.60	0.59	0.68	0.69	0.47	0.44
	0.67	0.68	0.67	0.68	0.76	0.79	1.0	1.0
Confetti	0.05	0.06	0.05	0.06	0.06	0.08	0.22	0.20
	0.45	0.43	0.44	0.41	0.47	0.45	0.76	0.74
Paper shreds	0.15	0.19	0.14	0.18	0.20	0.29	0.43	0.42
	0.60	0.62	0.59	0.60	0.61	0.65	0.84	0.77
PE granules	0.40	0.43	0.35	0.38	0.40	0.44	0.40	0.28
	0.62	0.64	0.59	0.64	0.64	0.71	0.84	0.88

mentioned above. Table III exhibits the evaluation of tracking performance with and without postprocessing. Since only the particle trajectories originating from the lance are of research interest for further analysis of the particles' properties and behaviors, we, therefore, labeled and evaluated merely these trajectories.

As shown in Table III, the tracking performances of both GNN and JPDA are relatively inaccurate because of the incompleteness and inexactitude of the tracklets. The presented postprocessing approach benefits a precise tracking performance, as indicated in Table III. While α and β of the direct tracking outcomes from GNN and JPDA are relatively low with values of roughly 0.2–0.3 for the majority of the fractions, the tracking results with postprocessing provide considerably higher α and β values with values around 0.5–0.6. JSC and JSC_θ are also significantly optimized with up to about 40%. Despite the significant improvement of the JSC by our postprocessing for confetti (0.06–0.47 and 0.08–0.45, respectively), the values for confetti remain below those for the other fuels, which can be attributed to their poor visibility due to their small sizes.

In general, GNN and JPDA provide comparable outcomes for tracking RDF fuel particles in the work. For measurements concerning position precision, for instance, α, β, and JSC, JPDA indicates a slight advantage of a few percent since GNN might assign several noises to tracks, which results in deviations of a few positions. Because JPDA considers all detections within the gate for the update, the deviation caused by noise is compensated. As for the aspect of track accuracy, the difference is even smaller than position precision. The most significant disparity occurs when tracking paper shreds, whose relatively large size and intricate gray value distribution lead to more FP detections that could result in several FP trajectories using JPDA.

Although GNN and JPDA can offer comparable performance, the computational cost and entailed corresponding storage capacity of JPDA are much higher, which disables its application for long-time tracking with a large number of detections. However, long-time tracking to accumulate sufficient complete trajectories serves as the prerequisite for analyzing particles' flight and combustion behaviors.

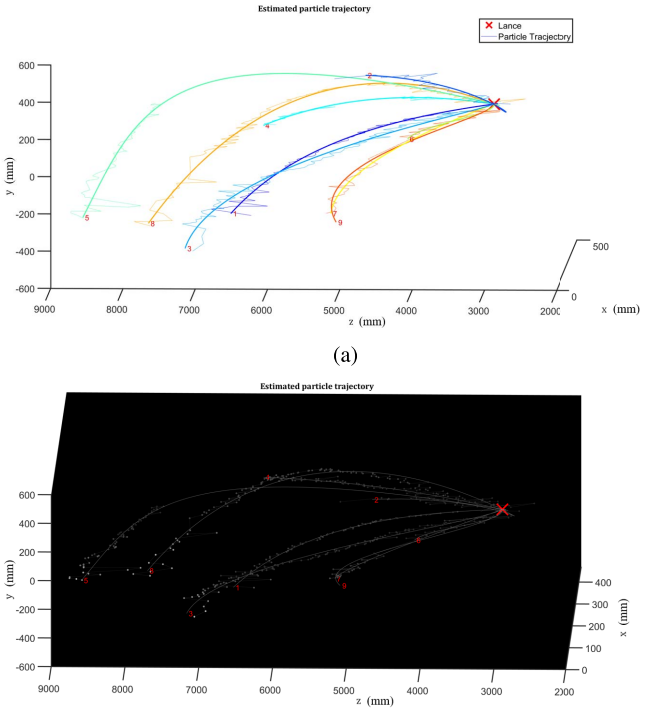


Fig. 18. Spatial trajectories of wood chips. (a) Estimated polynomials. Red Cross stands for the lance, thick solid lines are the estimated polynomials, and thin solid lines represent the direct converted 3-D trajectories. (b) Estimated polynomials with corresponding particle gray values. The red cross stands for the lance, thin solid lines are the estimated polynomials, and points represent the corresponding gray values of the detections.

C. 3-D Trajectory Estimation

With the completion of 2-D particle tracking, the trajectories are converted into 3-D in accordance with the 3-D coordinates provided by the camera. Here, we only extract the 3-D information for obtained trajectories from GNN for the following three reasons: 1) JPDA is not appropriate for long-time tracking with many detections; 2) JPDA computes the average coordinates of all detections within the gate, which might enhance the fluctuation of the converted 3-D trajectories; and 3) since the tracking performances are comparable, using GNN entails less computational cost and storage capacity. Due to the considerable fluctuations of the 3-D coordinates provided by the camera, we estimate the obtained 3-D trajectories with polynomials independently in three spatial directions on the condition that each trajectory originates from the lance using a least-squares approach. The directly converted 3-D trajectories together with their corresponding estimated polynomials are exampled in Fig. 18.

VI. CONCLUSION AND OUTLOOK

This article introduces approaches to detect and track RDF particles based on images captured by a high-speed plenoptic camera. First, a novel combined detection method that combines the detection results of a 2-D gray value-based detection approach and a 3-D-clustering method providing superior performance is presented. Subsequently, we utilize the 2.5-D linear Kalman filter together with GNN and JPDA,

respectively, to associate temporary particle detections for the sake of particle tracking.

To alleviate the issues occurring in the preliminary tracking results caused by detection inaccuracy, we developed a postprocessing framework, which takes full advantage of the motion similarity between tracklets to connect and merge tracklets to trajectories and fuse trajectories afterward. The proposed postprocessing method for multiple-particle tracking is applied to the tracking results of GNN and JPDA. The results obtained by the presented postprocessing framework significantly outperform the standard tracking methods.

Since, for performance reasons, JPDA is not appropriate to follow a significant number of detections for an extended period, the particles from the whole sequence with roughly 3000 frames are tracked barely by GNN. These acquired 2-D trajectories are converted into 3-D particle trajectories, which are estimated by polynomials to compensate for the fluctuations in depth measurement from the camera afterward. The estimated polynomials enable a statistical analysis of the flight behaviors of various RDF particles.

Our experiments with various fuel particles indicate the adequacy of the proposed tracking-by-detection process for the preliminary investigation of RDFs' flight properties. Future research will aim at investigating the adaptability and robustness of the presented approaches for the detection and tracking of other objects and developing an integrated tool for multiple-object tracking-by-detection. For multiple-particle detection, we focus on introducing the concept of combining 2-D gray value information and 3-D point clouds, and chiefly experiment with the combination of 2-D-SIFT and DBSCAN. Principally, in our combined particle detection framework, the applied 2-D detection approaches and 3-D clustering algorithms can easily be substituted, which is also a research topic in the future. In addition, we will further research the flight and combustion properties of the RDF particles.

REFERENCES

[1] M. Zhang et al., "Analysis of spatial flight properties of refuse-derived fuel particles based on a light-field camera system," in *Proc. 13th Eur. Conf. Ind. Furnaces Boilers* Algarve, Portugal, Dec. 2022, pp. 1–10.

[2] R. Streier et al., "Determination of the aerodynamic properties of refuse derived fuel by computer vision," in *Proc. 13th Eur. Conf. Ind. Furnaces Boilers*, Algarve, Portugal, 2022, pp. 135–150.

[3] M. N. Pedersen, M. Nielsen, S. Clausen, P. A. Jensen, L. S. Jensen, and K. Dam-Johansen, "Imaging of flames in cement kilns to study the influence of different fuel types," *Energy Fuels*, no. 31, pp. 11424–11439, Nov. 2017.

[4] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.

[5] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, 2001, pp. 456–567.

[6] M. Vogelbacher, "Eine neue Methode zur kamerabasierten analyse von Mehrstoffbrennern in industriellen Verbrennungsprozessen," Ph.D. thesis, Dept. Comput. Sci., Inst. Automat. Appl. Inform., Karlaruhe Institute of Technology, Karlsruhe, Germany, 2018.

[7] J. Matthes, J. Hock, P. Waibel, A. Scherrmann, H.-J. Gehrmann, and H. B. Keller, "A high-speed camera based approach for the on-line analysis of particles in multi-fuel burner flames," *Experim. Thermal Fluid Sci.*, vol. 73, pp. 10–17, May 2016.

[8] Y. Xuan, P. Jihong, and Y. Wanhai, "Firing particle flow detection and tracking in sequence images," in *Proc. 3rd World Congr. Intell. Control Autom.*, Jul. 2000, pp. 2666–2670.

[9] W. V. Nicholson and R. M. Glaeser, "Review: Automatic particle detection in electron microscopy," *J. Struct. Biol.*, vol. 133, nos. 2–3, pp. 90–101, Feb. 2001.

[10] P. C. Miller and R. S. Caprari, "Demonstration of improved automatic target-recognition performance by momentanalysis of correlation peaks," *Appl. Opt.*, vol. 38, pp. 1325–1331, May 1999.

[11] K. R. Lata, P. Penczek, and J. Frank, "Automatic particlepicking from electron micrographs," *Ultramicroscopy*, vol. 58, pp. 381–391, Aug. 1995.

[12] J. J. Merelo, A. Prieto, F. Moran, R. Marabini, and J. M. Carazo, "Automatic classification of biological particles fromelectron-microscopy images using conventional and genetic-algorithm optimized learning vector quantization," *Neural Process. Lett.*, vol. 8, pp. 55–65, Aug. 1998.

[13] H. Irshad, A. Veillard, L. Roux, and D. Racoceanu, "Methods for nuclei detection, segmentation, and classification in digital histopathology: A review—Current status and future potential," *IEEE Rev. Biomed. Eng.*, vol. 7, pp. 97–114, 2014.

[14] P. D. Wellner, "Adaptive thresholding for the DigitalDesk," Xerox, EuroPARC Tech. Rep. EPC-93-110, 1993.

[15] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, Jan. 1988.

[16] J. MacQueen et al., "Some methods for classifcation and analysis of multivariate observations," in *Proc. 55th Berkeley Symp. Math. Statist. Probab.*, Boston, MA, USA, 1967, pp. 156–189.

[17] H. Liu, J. Luo, P. Wu, S. Xie, and H. Li, "People detection and tracking using RGB-D cameras for mobile robots," *Int. J. Adv. Robotic Syst.*, vol. 13, no. 5, Sep. 2016, Art. no. 172988141665774.

[18] Y.-H. Chiang, C.-M. Hsu, and A. Tsai, "Fast multi-resolution spatial clustering for 3D point cloud data," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2019, pp. 1678–1683.

[19] M. Ester, J. Sander, H.-P. Kriegel, and X. Xu, "DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN," *ACM Trans. Database Syst.*, vol. 42, no. 3, pp. 1–21, Aug. 2017.

[20] J. Huang, Q. Zhu, L. Yang, D. Cheng, and Q. Wu, "QCC: A novel clustering algorithm based on quasi-cluster centers," *Mach. Learn.*, vol. 106, no. 3, pp. 337–357, 2017.

[21] Z. Wang, "A new clustering method based on morphological operations," 2019, *arXiv:1905.10548*.

[22] I. Halima, J.-M. Laferté, G. Cormier, A.-J. Fougères, and J.-L. Dillenseger, "Depth and thermal information fusion for head tracking using particle filter in a fall detection context," *Integr. Comput.-Aided Eng.*, vol. 27, no. 2, pp. 195–208, Feb. 2020, doi: 10.3233/ICA-190615.

[23] M. Attamimi, T. Nagai, and D. Purwanto, "Object detection based on particle filter and integration of multiple features," in *Proc. INNS Conf. Big Data Deep Learn.*, 2018, pp. 17–19.

[24] D. Justen, "Untersuchung eines neuartigen 2D- gestützten 3D-PMD-bildverarbeitungssystems," Ph.D. thesis, Univ. Siegen, Siegen, Germany, 2001.

[25] T. Ophoff, K. V. Beeck, and T. Goedemé, "Exploring RGB+depth fusion for real-time object detection," *Sensors*, vol. 19, no. 4, p. 866, 2019.

[26] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME, D, J. Basic Eng.*, vol. 82, pp. 35–45, Oct. 1960.

[27] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[28] P. Konstantinova, A. Udvarev, and T. Semerdjiev, "A study of a target tracking algorithm using global nearest neighbor approach," in *Proc. 4th Int. Conf. Conf. Comput. Syst. Technol.*, 2003.

[29] Y. Bar-Shalom and E. Tse, "Tracking in a cluttered environment with probabilistic data association," *Automatica*, vol. 11, no. 5, pp. 451–460, 1975.

[30] T. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Multi-target tracking using joint probabilistic data association," in *Proc. 19th IEEE Conf. Decis. Control Including Symp. Adapt. Processes*, Dec. 1980, pp. 807–812.

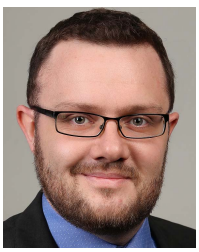
[31] A. Milan, S. H. Rezatofighi, A. Dick, I. Reid, and K. Schindler, "Online multi-target tracking using recurrent neural networks," in *Proc. 31st AAAI Conf. Artif. Intell. (AAAI)*, San Francisco, CA, USA, 2017, pp. 190–230.

[32] M. Ahmad, I. Ahmed, F. A. Khan, F. Qayum, and H. Aljuaid, "Convolutional neural network–based person tracking using overhead views," *Int. J. Distrib. Sensor Netw.*, vol. 16, no. 6, Jun. 2020, Art. no. 155014772093473.

[33] D. Sun, X. Wang, Y. Lin, T. Yang, and S. Wu, "Introducing depth information into generative target tracking," *Frontiers Neurorobot.*, vol. 15, pp. 1–5, Sep. 2021.
[34] C. Ma et al., "Trajectory factory: Tracklet cleaving and re-connection by deep Siamese Bi-GRU for multiple object tracking," in *Proc. IEEE Int. Conf. Multimedia Expo. (ICME)*, Nov. 2018, pp. 1–6.
[35] J. Peng et al., "TPM: Multiple object tracking with tracklet-plane matching," *Pattern Recognit.*, vol. 107, Nov. 2020, Art. no. 107480.
[36] M. Zhang, J. Matthes, K. Aleksandrov, H.-J. Gehrmann, and M. Vogelbacher, "Detection and tracking of refuse-derived fuel particles based on a light field camera system," *Technisches Messen*, vol. 88, no. 6, pp. 396–407, 2021.
[37] K. Aleksandrov et al., "Waste incineration of Polytetrafluoroethylene (PTFE) to evaluate potential formation of per- and poly-fluorinated alkyl substances (PFAS) in flue gas," *Chemosphere*, vol. 226, pp. 898–906, Dec. 2019.
[38] Y. Zhang and T. Funkhouser, "Deep depth completion of a single RGB-D image," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 1–4.
[39] Zhang, Huang, and Zhao, "A new model of RGB-D camera calibration based on 3D control field," *Sensors*, vol. 19, no. 23, p. 5082, Nov. 2019.
[40] H. Sandemann, "Untersuchungen zur Genauigkeit von Streckenmessungen mit einer fokussierten plenoptischen Kamera im Fernbereich," in *Proc. Photogrammetrie Laserscanning Optische 3D-Messtechnik, Beiträge der Oldenburger 3D-Tage*, 2016.
[41] C. Perwaß and L. Wietzke, "Single lens 3D-camera with extended depth-of-field," in *Proc. Int. Soc. Opt. Eng. (SPIE)*, vol. 8291, 2012, p. 4.
[42] M. Zhang, J. Matthes, K. Aleksandrov, H.-J. Gehrmann, and M. Vogelbacher, "Light-field camera based spatial multiple small particle tracking with post-processing," in *Advances in Automation, Mechanical and Design Engineering*. Cham, Switzerland: Springer, 2022, pp. 95–123.
[43] J. Munkres, "Algorithms for the assignment and transportation problems," *J. Soc. Ind. Appl. Math.*, vol. 5, no. 1, pp. 32–38, 1957.
[44] K. I. Laws, "Texture energy measures," in *Proc. Image Understand. Workshop*, 1979, pp. 47–51.
[45] S. Dash and U. R. Jena, "Multi-resolution laws' masks based texture classification," *J. Appl. Res. Technol.*, vol. 15, no. 6, pp. 571–582, Dec. 2017.
[46] N. Chenouard et al., "Objective comparison of particle tracking methods," *Nature Methods*, vol. 11, pp. 281–289, Jan. 2014.

Miao Zhang received the B.Sc. degree in vehicle engineering from Tongji University, Shanghai, China, in 2015, and the M.Sc. degree in mechanical engineering from the Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2018, where she is currently pursuing the Ph.D. degree in mechanical engineering with the Faculty of Mechanical Engineering.

Since 2019, she has been a Ph.D. Researcher with the research group "Advanced Automation Methods for Industrial Processes," Institute for Automation and Applied Informatics, KIT. Her research focuses on image processing based on light field camera systems.



Markus Vogelbacher received the Dr.Ing. degree in mechanical engineering from the Department of Mechanical Engineering, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2018.

Since 2015, he has been a Scientist with the working group "Advanced Automation Methods for Industrial Processes," Institute for Automation and Applied Informatics, KIT. His research focuses on image processing and machine learning methods.

Dr. Vogelbacher received the NAMUR Award 2018 for his dissertation.

Veit Hagenmeyer was born in Germany in 1971. He received the Ph.D. degree from the Laboratoire des Signaux et Systèmes (L2S), C.N.R.S.-Supélec, Université Paris-Sud, Bures-sur-Yvette, France, in 2002.

From 2002 to 2003, he held a post-doctoral position at L2S, C.N.R.S.-Supélec, Université Paris-Sud. From 2003 to 2006, he was a Research Engineer with the Advanced Process Control Group, BASF SE, Ludwigshafen, Germany. From 2007 to 2008, he was the Head of the Advanced Process Control Group, BASF SE. From 2008 to 2009, he was a Senior Consultant for Verbund Simulation of BASF sites in Europe. From 2009 to 2010, he was the Personal Assistant to the European Site Manager Europe (Level 2 of BASF Group). From 2010 to 2014, he was the Power Plant Director of BASF (three power plants and energy grids). In 2014, he became the Director of the Institute of Automation and Applied Informatics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, where he became a Full Professor of energy informatics at the Faculty of Informatics.

Krasimir Aleksandrov studied chemical engineering at the University of Chemical Technology and Metallurgy (UCTM), Sofia, Bulgaria, from 1994 to 1999. He completed his doctoral thesis in inorganic solid-state chemistry at the University of Siegen, Siegen, Germany.

At the Dechema Research Institute, Frankfurt am Main, Germany, he supervised research projects on the development of novel corrosion protection coatings for high-temperature applications as a Research Associate. During his stay at the Karlsruhe University of Applied Sciences, Karlsruhe, Germany, he developed filtration systems for drinking water treatment. Since 2018, he has been a Staff Member with the Combustion Technology Group, Institute of Technical Chemistry, Karlsruhe Institute of Technology (KIT), Karlsruhe. His research focuses on the combustion of waste in the grate and rotary kiln furnaces, as well as the development of a fuel cell for highly efficient direct conversion of renewable carbon into electrical energy.

Hans-Joachim Gehrmann initially studied mechanical and process engineering at the University of Stuttgart, Stuttgart, Germany, and the Technical University of Clausthal, Clausthal-Zellerfeld, Germany, from 1988 to 1995, before completing his doctorate on the subject of "Pyrolysis of waste in rotary kiln systems" as part of his employment at the Clausthaler Umwelttechnik Institut GmbH (CUTEC), Clausthal-Zellerfeld. He is currently pursuing the Ph.D. degree with the Chair "Processes and Environment" of Prof. Beckmann, Bauhaus-Universität Weimar, Weimar, Germany.

At the CUTEC Institute, he supervised projects on application-oriented research with industrial partners on thermochemical processes of pyrolysis, gasification, and incineration for biomasses and wastes as a Research Associate. After completing his doctorate, he moved to the Karlsruhe Research Center (now KIT), Karlsruhe, Germany, in 2006, as a Group Leader, and is responsible for scientific issues relating to the combustion of waste in the grate and rotary kiln furnaces. Together with the company W.L. Gore & Associates, Inc., Elkton, MD, USA, for example, a study was conducted on the fate of fluorine-containing substances during incineration. When he took over as head of the department in 2018, he expanded his field of work to include particle technology; current research projects include the characterization of carbon fibers from the dismantling of wind turbines.

Jörg Matthes studied mechatronics at the TU Bergakademie Freiberg, Freiberg, Germany. He received the Dr.Ing. degree from the Faculty of Mechanical Engineering, TH Karlsruhe, Karlsruhe, Germany, in 2004.

In 2017, the Karlsruhe Institute of Technology (KIT), Karlsruhe, awarded him the teaching license for the subject "machine vision in automation technology." Since 2000, he has been a Scientist with the Institute for Automation and Applied Informatics, KIT, where he heads the working group "Advanced Automation Methods for Industrial Processes." His research focuses on image processing and process control.