Bulk cargo volume measurement for moving dump trucks with a single – layer LiDAR and a camera

D.A. Bocharov [1,2], V.V. Kokhan [1], I. D. Konyushenko [1,3], A. Y. Resniansky [2], I.P. Nikolaev [1,2], D.P. Nikolaev [2,4]
[1] Institute for Information Transmission Problems, 127051, Russia, Moscow, Bolshoy Karetny per. 19, build. 1;
[2] Federal Research Center "Computer Science and Control", Russian Academy of Sciences, 119333, Russia, Moscow, Vavilova 44, build. 2;
[3] Moscow Institute of Physics and Technology, 141701, Russia, Dolgoprudny, Institutskiy per. 9;
[4] Smart Engines Service LLC, 117312, Russia, Moscow, pr. 60–letiya Oktyabrya 9

Abstract

The paper addresses the problem of non–contact bulk cargo volume estimation for moving dump trucks. A common scanning method that lets to evaluate the volume of cargo of complex surface for a moving truck implies two single–layer (2D) Light Detection and Ranging (LiDAR) sensors: one is used to scan a vehicle in a plane perpendicular to its movement and the second — to estimate vehicle displacements and restore scans positions on an axis along vehicle movement direction. While LiDAR sensors provide reliable measurement signals in controlled environments their efficacy drastically decreases under challenging outdoor conditions: sand dust, fog, rain heavy precipitation cause false detections and distort LiDARs signal. Thus, vehicle displacements estimated with a highly corrupted LiDAR signal can not be used for a reliable measurement as they may lead to significant volume calculation errors. Partially this is solved in multi–echo lidar where distorted data could be separated from the relevant. In contrast to the single–echo 2D LiDAR, image data from industrial cameras is less sensitive to sand dust or fog. In the paper we propose a novel bulk cargo estimation method that implies only one 2D LiDAR and for vehicle displacements estimation utilizes a camera and computer vision methods. As we demonstrate on a diverse dataset of 730 pairs of dump truck passes from an operating sand pit, the proposed method is more accurate than the two 2D LiDARs baseline while requiring a significantly cheaper sensor. In case if a camera is already present in the volume measurement system and utilized for loaded material classification then the proposed method lets to reduce the cost of solution by the cost of one lidar.

Keywords: bulk cargo, dump trucks, volume measurement, LiDAR, camera, optical flow, deep learning, outdoor, sand.

Citation: Bocharov DA, Kokhan VV, Konyushenko ID, Resniansky AY, Nikolaev IP, Nikolaev DP. Bulk cargo volume measurement for moving dump trucks with a single–layer LiDAR and a camera. Computer Optics 2025; 49(6): 1120-1128. DOI: 10.18287/COJ1801.

Introduction

Automated systems for bulk materials volume measurement are used as a much more precise alternative to manual accounting of volumes of transported materials and are utilized in mining, development, agriculture and other industries. A commonly used practice is based on volume estimation through measured weights with known material type, its density, humidity etc. While weight measurements can be easily obtained using industrial scales, inaccurate estimation of any of the material parameters can crucially affect the volume estimation. Stable truck overloads in addition to financial losses for material suppliers affect road quality, increase risks of traffic accidents and infrastructure damage [1]. Another approach for load estimation implies Light Detection And Ranging (LiDAR) sensors to obtain point cloud measurements directly in the metric coordinate system from which the volume information is retrieved [2 – 4].

Existing LiDAR–based cargo volume measurement systems for dump trucks perform vehicle scanning in static or dynamic modes, i.e. with or without requirement for a vehicle to stop during scanning. Dynamic mode compared to static brings a significant evaluation speed–up that is crucial for high load sites with frequent loading and unloading operations and therefore becomes a highly valuable advantage.

In order to obtain the carried bulk material volume for a moving truck two single–layer (2D) LiDARs is enough [5]. One is used to scan a vehicle in a plane perpendicular to its movement, thus obtaining a layer–by–layer scan of the entire truck. The second is mounted to scan towards vehicle movement and is used to estimate its displacements and locate scans along the direction of vehicle movement (LiDAR+LiDAR setup, see Fig. 1*a*). After 3D point clouds for an empty and loaded truck are reconstructed the bulk cargo volume can be calculated by subtracting the volume of an empty truck from that of loaded.

On the one hand LiDARs are proved to be reliable and precise distance measurement sensors, but on the other they are highly expensive equipment that constitutes a significant part of the entire hardware system cost [6]. We propose a new optimized LiDAR+camera setup that preserves one LiDAR for transverse truck scanning and for velocity estimation utilizes a monocular video camera that is much more budget friendly than the 2D LiDAR. With the help of the remaining LiDAR for transverse scanning, pixel shifts retrieved from a calibrated camera can be accurately transformed into metric units.

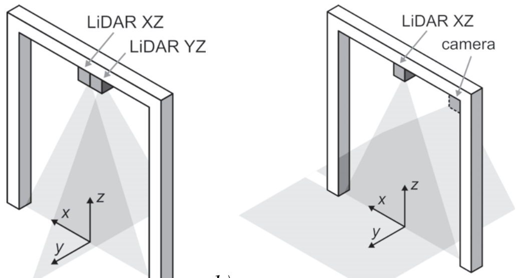

Fig. 1. Schemes for the common cargo measurement setup with two 2D LiDARs (a) and for the proposed configuration with one LiDAR and a camera that is used for vehicle velocity estimation (b)

Besides better affordability, cameras possess other important advantages compared to LiDARs. Firstly, a camera provides valuable data for general monitoring and being mounted above a vehicle can also be used for automatic cargo type classification (sand, gravel, material fractions, household waste). Secondly, image data quality is less affected by rain, snow, sand dust or fog and even in cases of noticeable degradation under severe observation conditions its quality may be enhanced with the help of image processing methods. Since cargo volume measurement systems generally operate outdoors, LiDARs usage has a risk of signal corruption in challenging outdoor conditions: particles in space between a LiDAR and an object may produce a false detection in case the reflected signal level is high enough to be considered as relevant [7].

The LiDAR+camera setup still preserves one single–echo 2D LiDAR and thus is not completely free from data corruption risks if snow or sand dust is present. However the data from the transverse scanning LiDAR XZ has more potential for recovery in cases of non–extreme conditions compared to that of LiDAR YZ. LiDAR XZ scans the passing truck layer–by–layer, thus the entire scanning contains its spatial information which can help to detect and filter clutter in the signal.

In this study we investigate the task of vision–based truck displacement estimation using a calibrated monocular camera. We decompose this task into two subproblems: vehicle pixel displacement estimation and their conversion into the metric units. In this paper we propose to use the NeuFlow–v2 neural network model by Z. Zhang et. al for optical flow estimation [8] and propose an algorithm for units conversion that implies depth values from the remaining LiDAR for transverse scanning. We demonstrate that the proposed LiDAR+camera method overcomes the accuracy of the LiDAR+LiDAR setup while requiring one less expensive sensor.

The main contributions of the paper are the following:

1. To the best of our knowledge we are the first to present a method for bulk cargo volume measurement for moving dump trucks using a single 2D LiDAR and a camera;
2. We demonstrate that the LiDAR+camera method provides more accurate volume estimation results in outdoor environments than the state–of–the–art LiDAR+LiDAR approach;
3. A new algorithm for object metric velocity estimation using a camera and 2D LiDAR is proposed.

1. Related work

The related studies cover different cargo volume evaluation tasks: precise volume estimation, empty or full truck body classification, loading fraction classification, etc. Therefore existing scanning methods imply different technical approaches and sensors depending on the considered problems.

LiDAR–based methods for volume estimation are presented in papers [1, 2, 5, 9]. In [1] authors use a multi–layer (3D) LiDAR mounted on an excavator's boom and oriented towards a truck body. The paper [2] presents a method for cargo volume estimation for moving trucks using a pair of multi–layer LiDARs. E. Duff in his study [5] describes a method that implies two 2D LiDARs mounted above a passing truck with orthogonal scanning planes towards and perpendicular to trucks movement. A relevant LiDAR–based approach for bulk cargo volume measurement for moving vehicles is described in the patent [10]. The invention is similar to the method by E. Duff which utilizes two orthogonally mounted 2D LiDARs above the passing dump truck one of which is used for vehicle displacements estimation. In [9] authors propose a subtraction–based approach that utilizes estimated linear sizes of cargo, thus is not applicable for bulk cargo with a complex form.

Another approach for volume estimation is vision–based. Some works consider volume measurement as a loaded truck classification task without 3D scene reconstruction and analysis. In [11] vision–based methods are investigated for a binary task of empty and full–load truck classification. The authors evaluate different deep–learning models and demonstrate VGG16 model overall superiority among all considered models. In [12] Sun X. et al. investigate the task of loaded trucks classification into categories defining truck load percentage. Authors proposed a method based on the VGG16 model with additional regression stage for volume value estimation. On real data of loaded mining trucks their method demonstrates an average absolute error of 2.5 m^3. For method evaluation data of only a single vehicle were used. Similar to [12] VGG16–based method is proposed in [13]. In contrast to works mentioned above in [14] authors propose a solution of volume estimation using a single depth camera that is mounted on an excavator's boom like in [1]. In [15] authors propose a vision–based construction waste volume estimation method using a single monocular camera. With geometrically calibrated data authors propose a cargo objects reconstruction algorithm from which volume information is retrieved. The proposed method explicitly utilizes the information about cubic form of truck bucket bodies which limitates its applicability for truck models with more complex body forms.

In this paper we propose a new hybrid LiDAR+camera method that utilizes only one 2D LiDAR and a videocamera. In contrast to prior work the proposed method preserves advantages of LiDAR+LiDAR methods such as volume evaluation for a complex cargo form and dynamic measurement mode, but requires 1 expensive sensor less. The measurement point cloud for a moving vehicle is obtained by a single–layer LiDAR and the restoration of scans locations along trucks movement is performed using a video camera.

2. Sensing setup

Sensors are mounted on a rectangular frame with 6 meters width and 6.5 meters height (see Fig. 1) located on a rural road by the bulk cargo loading site. In the middle of the crossbar an 2D LiDAR is mounted so that the scanning plane is perpendicular to trucks movement along the road (further we will denote it as LiDAR XZ). The visibility angle range of the LiDAR XZ covers the entire road surface. In order to locate XZ scans on the Y axis we propose to use a camera mounted in a top corner of the frame. The optical axis of the camera is approximately orthogonal to the vehicle movement direction. The extrinsic parameters of the camera and the LiDAR XZ are controlled during the set up and are therefore known. The projection of XZ LiDAR' sensing plane on the camera sensor is considered to be a central vertical line on an image. Since the LiDAR XZ is mounted so its sensing plane lies within the plane of the frame, the location of the visible part of the vertical column (see Fig. 2) of the frame is used to check the camera extrinsic parameters.

We use a monochrome camera in 2×2 binning mode that provides uncompressed video in grayscale with resolution 1024 × 768 pixels. Since we investigate the task of moving objects analysis we utilize cameras with global shutter sensors. The gain and exposure parameters are set in ranges from 20 to 24 and from 20 to 7500 mcs correspondingly and are controlled automatically by the camera driver. With described parameters video frame rate is approximately 18 frames per second. The focus length is 8 mm with field of view approximately 48°×34° (H×W). Four near infrared spotlights mounted in top corners of the frame are used for lighting during the dark hours. IR filter is applied to equalize day and night lighting conditions for camera without SWIR lens and day and night filter switch. Frame examples of analyzed videodata highlighting the diversity of observation conditions are shown in Fig. 2. The description of the utilized equipment is given in Tab. 1 below.

Fig. 2. Examples of input data used for truck movement analysis in different conditions (from left to right, top to bottom): soft lighting, illuminated by the IR spotlights, overexposure, snow, shadows and sand dust

Tab. 1. Description of the utilized equipment

Equipment	Parameters
LiDAR	Gtek LSD102A with angular resolution 0.25° at scanning frequency 50 Hz. Repeatability ± 1 cm
Camera	Daheng Imaging MER2–302–37GM with 1/1.8 inch sensor and resolution 2048×1536 pixels
Lens	Daheng Imaging HN–P–0828–6M–C1/1.8 with angle of view (H×W) 48.58°×33.84° and infrared filter at 800 nm
Illumination	4 Infrared light spots at 850nm with radiation angle 52°. Radiation power 6W/sr.

3. Volume measurement method with a LiDAR and a videocamera

We propose a hybrid LiDAR+camera approach for bulk cargo measurement for moving trucks that implies a single–layer LiDAR for truck transverse scanning and a monocular camera for displacements estimation. With synchronized layer–by–layer scannings of a truck and obtained displacements in meter units the 3D point cloud of the observed truck surface is restored. Finally, the goal bulk cargo volume is obtained as a difference between volumes of air above the truck body for a pair of measurements corresponding to an empty and loaded passes of the same truck.

The proposed object displacement estimation algorithm with a 2D LiDAR and a camera can be described as follows. With two given consecutive frames the optical flow map within the region of interest corresponding to the LiDAR XZ scanning plane is estimated. The LiDAR XZ provides depth information that is used to estimate a map of distances from the sensor plane of the calibrated camera spatially corresponding to the flow map. At the final stage for given flow and distance maps object displacement in meter units is estimated using a preliminary trained linear model.

3.1. Pixel displacement estimation

For pixel displacement retrieval we propose to use optical flow methods. Comparing dense and sparse optical flow, dense is more preferable: sparse methods can not guarantee estimations to belong to the region of interest. Nevertheless sparse estimations may be followed by additional interpolation to obtain values in the required locations but as it will be shown in the experiments section the dense solution provides more accurate results.

A common dense optical flow baseline refers to the Lucas–Kanade algorithm [16]. For our task the Lucas–Kanade algorithm is not preferable since it is not robust to challenging outdoor environments and local motion non–uniformities due to static background, occlusions and shadows. For dense optical flow we utilize a NeuFlow–v2 neural network model proposed by Z. Zhang et al. [8], which in 2024 demonstrated state–of–the–art performance while being significantly more computationally effective than the competitors.

Input images to NeuFlow–v2 were downscaled to 432×768 and the model output image was restored to 1024×768. The flow values used for displacement evaluation are sought within the XZ plane projection on an image that is approximated with a vertical line (see Fig. 4). On that stage the pixel displacements corresponding to the projected XZ plane for the *t*–th timestamp $D_{pix,t} = [d_{pix,t,1}, \ldots, d_{pix,t,n}]$ are returned.

In fact a single $d_{pix,t}$ taken from a pixel corresponding to a truck is enough to get the sought displacement, but for better reliability of the estimate we analyze a set of n displacement values. In the current work $n = 15$, on the XZ projection points locations are distributed uniformly.

3.2. Depth estimation

Since the observed truck surface is complex and non–planar, the observed velocities of truck pixels sufficiently depend on the depth from the camera to the corresponding point on a truck. We propose an algorithm for depth map calculation corresponding to the evaluated pixel displacement value.

Since we have only a monocular camera we fuse camera and LiDAR data, assuming the presence of time annotation for each shot, to get an accurate depth estimation. Given an image, we use a synchronized LiDAR slice to calculate a 2D–polygon of a truck slice. After that we estimate distances to the truck points in the plane of the setup frame. By that means for a given set of pixel displacements ${D^{pix}}_t = [{d^{pix}}_{t,1}, \ldots, {d^{pix}}_{t,n}]$we obtain $L_t = [l_{t,1}, \ldots, l_{t,n}]$ that corresponds to a number of estimated distances to the object surface.

3.3. Displacement estimation in meter units

We assume a pinhole camera model with known focal length f and pixel size s in millimeters (see Fig. 3). To convert a pixel shift ${d^{pix}}_{t,i}$ into meter units ${d^{mm}}_{t,i}$ we imply the following formula:

$${d^{mm}}_{t,i} = {d^{pix}}_{t,i}\, l_{t,i}\, s\, f^{-1},$$

where $l_{t,i}$ is an estimated distance to a truck point in millimeters. We assume the geometric distortion to be insufficient.

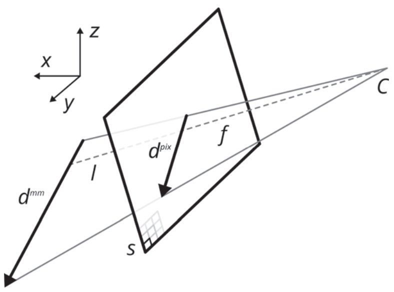

Fig. 3. The pinhole camera model, where C is the center of the pinhole camera, f denotes the focal length, d^{pix} and d^{mm} are displacements in pixels and millimeters correspondingly, l is the estimated distance to an object and s is the linear size of a square pixel

With calculated ${d^{mm}}_{t,i}$ the final displacement in meter units is obtained as a median value that provides an estimate robust to possible outlying values:

$${d^{mm}}_t = median_i\, {d^{mm}}_{t,i}.$$

3.4. Fine–tuning

In order to fine–tune the displacement estimation algorithm we imply a linear model parameterized with an α parameter: ${\hat{d}^{mm}}_t = \alpha\, {d^{mm}}_t$. As a loss function for linear model optimization a mean absolute volume estimation error was utilized. For optimization the binary search algorithm was used.

3.5. Algorithm data–flow diagram

Fig. 4 presents the data–flow diagram of the proposed displacements estimation algorithm using a video camera and a LiDAR XZ.

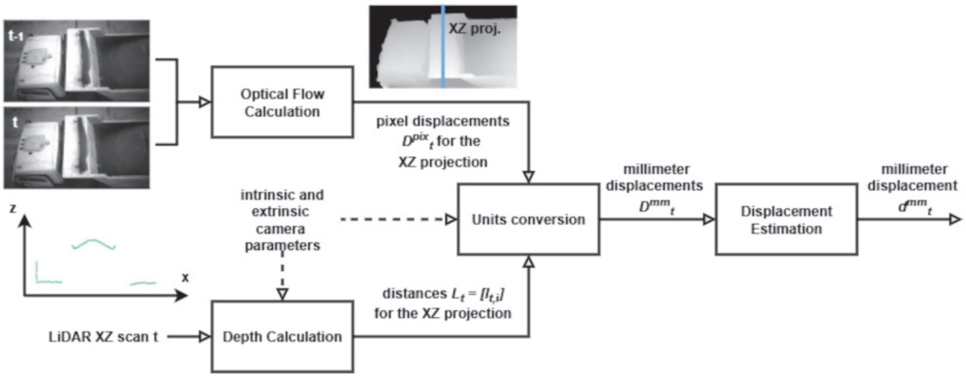

Fig. 4. Data–flow diagram of the proposed displacement estimation algorithm

3.6. Volume measurement

The volume measurement method is based on the subtraction approach according to which the target volume value is obtained as a difference between measured volume of air above the truck bodies at departure and arrival. The input data for each pass consists of LiDAR XZ scans $S_{pass} = [s_{pass,1}, \dots, s_{pass,n}]$ and estimated millimeter displacements $\widehat{D}^{mm}{}_{pass} = [\hat{d}^{mm}{}_{pass,1}, \dots, \hat{d}^{mm}{}_{pass,n}]$, where $pass$ index refers to arrival or departure and n is the number of synchronized LiDAR scans and estimated displacements within the time interval of a truck passing under the sensing setup.

Let us consider the algorithm of air volume measurement for a single truck pass. It can be described with 3 main stages. On the first stage truck body on the LiDAR data is segmented [7]. The input LiDAR signals are represented with a 2D image, rows of which correspond to consecutive scans, columns correspond to scanning angles and the image intensity – to the measured distance in millimeters. The result of body segmentation is a pair of indices $[t_{front}, t_{rear}]$ that localize front and rear transverse sides and a list of coordinates $J_{left} = [(j,t)_{left,i}]$ and $J_{right} = [(j,t)_{right,i}]$ that localize left and right longitudinal sides correspondingly (see Fig. 5). On the second stage the body's height above the earth surface is estimated. That value is further utilized in order to consider the loaded truck suspension sag in the sought volume calculation stage under the assumption that the sag is uniform. On the last stage the sought air volume is obtained as the following sum: $V_{pass} = \sum_{i=t_{front}}^{t_{rear}} A_{pass,i}\ \hat{d}^{mm}{}_{pass,i}$, where $A = \sum_{j=x_{left}}^{x_{right}} z_j$ is the calculated area above the truck's body, x_{left}, x_{right} denote to coordinates of left and right longitudinal sides of body in XYZ coordinate system and z_j is the distance from the LiDAR XZ plane to the truck's body.

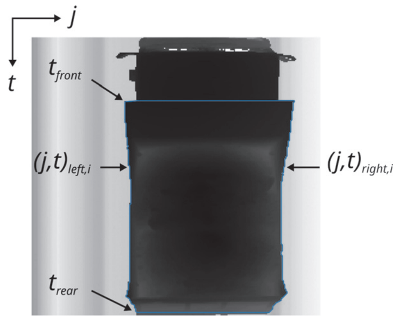

Fig. 5. Rasterized data from the LiDAR XZ for a truck pass with body borders segmentation results, where j denotes the scanning angle and t refers to the scan index

After $V_{arrival}$ and $V_{departure}$ are calculated the sought bulk cargo volume is obtained as a following difference: $V = V_{departure} - V_{sag} - V_{arrival}$, where V_{sag} denotes the volume that compensates the suspension sag of the loaded truck.

4. Experiments

4.1. Experiments methodology

In order to evaluate accuracy of object displacement estimation method different approaches may be utilized: a) accuracy estimation in pixel units on a video, b) accuracy of meter displacement estimation and c) evaluate the quality of cargo volume measurement system in general. Let's discuss these approaches.

The first is to estimate truck displacement on a video that lets us compare different vision–based methods. The ground–truth data may be obtained through manual keypoints tracking and estimating object shifts in pixel units. The main limitation of such an approach is inaccuracy of manual keypoints tracking that may affect the ground–truth and the fact that this approach excludes LiDAR–based from the competitors.

Quality of vehicle displacement estimation in metric units lets us evaluate the isolated modules used for LiDAR XZ scans positioning on the Y axis. For that purpose ground–truth displacements are essential. In such an approach estimations obtained with a LiDAR YZ may be considered as ground–truth. But as any measurement equipment it is prone to measurement errors under challenging conditions and the ground–truth also may occur distorted.

The first two approaches let us evaluate the quality of methods for object displacement estimation while the main objective is the accuracy of bulk cargo volume measurement. Thus, in this paper we propose to evaluate the volume estimation quality depending on the utilized sensors and methods. In this case the ground–truth information is provided in cubic meters by the material supplier, where for bulk material loading a calibrated excavator bucket with known volume is used. Since such a method may also suffer from inaccurate excavator bucket calibration, analysis of objects with precisely known volume may be a better alternative. Such investigation remains to be the future work.

The time interval within which the vehicle shifts are analyzed are determined by the pass detector module and are identical for all competing algorithms.

4.2. Datasets

Experiments were conducted using datasets of paired forward and backward truck passes before (empty) and after loading (loaded). For each pass the following data is available: LiDAR XZ and YZ data and a video. Input data for LiDAR+LiDAR setup is a pair of LiDAR XZ and YZ scans and for the LiDAR+camera — LiDAR XZ signal and a video. For each pass a ground–truth bulk cargo volume value is given.

The characteristics of the trucks velocities for empty (loaded) passes are the following. mean: 9.27 km / h (– 8.35 km/h), median: 9.29 km / h (– 8.28 km / h), 5th percentile: 5.22 km / h (– 4.63 km / h), 95th percentile: 12.92 km / h (– 12.68 km / h). Since the datasets do not possess displacement or velocity information the estimates above were obtained using the proposed NeuFlow–v2–based algorithm.

We utilized 2 non–intersecting datasets for training and tests. The train dataset was utilized for algorithms fine–tuning. The test dataset contains passes under diverse observation conditions that include clear, sand dust, overexposure and strong shadows. The passes in the test dataset were annotated with the following categories: “clear”, “dust”, “overexposure” and “shadow” in such a manner that each pass contained only one corresponding distortion. If there were one and more distortion types, for example sand dust and overexposure, it was annotated with the label “distored”. The volumes of the train and test datasets are 15 and 730 pairs of truck passes (totally 30 and 1460 passes correspondingly).

4.3. Metrics

To evaluate volume measurement quality we utilize mean (ME) and mean absolute errors (MAE) in absolute and relative units:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(|V_i - GT_i|),\ MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|V_i - GT_i|}{GT_i}\cdot 100\%,$$
$$ME = \frac{1}{n}\sum_{i=1}^{n}(V_i - GT_i),\ MPE = \frac{1}{n}\sum_{i=1}^{n}\frac{V_i - GT_i}{GT_i}\cdot 100\%.$$

where V_i is an estimation for *i*–th dataset item and GT_i is the corresponding ground–truth.

4.4. Vision–based competitors

We compare different modifications of the LiDAR+camera method with various algorithms for vehicle pixel displacement estimation. As for competitors we consider feature point–based algorithms commonly used for object tracking and movement analysis. Since feature point–based algorithms provide sparse optical flow the pipeline given in Fig. 4 was adapted in order to obtain flow values for the regions of interest corresponding to the XZ plane projection on an image. For that purpose we additionally applied aggregation of the estimated pixel shifts and distances using median filtration to obtain a scalar estimation instead of forming a set of corresponding estimations.

We consider 3 LiDAR+camera modifications that are based on sparse flow estimation with SIFT [17], ORB [18] feature points and ECO–TR [19] — the state–of–the–art neural network model for end–to–end feature points detection and matching. The algorithms were implemented using Python and OpenCV library.

4.5. LiDAR–based algorithm

As the baseline we consider an algorithm that processes signals from the LiDAR YZ, which senses the truck along its movement. Since the input LiDAR signals are already in meter units, the estimated displacement represents the truck shift in millimeters.

The iterative closest point (ICP) algorithm is a well known algorithm used for point cloud registration proposed in [5, 9] for truck displacement estimation by 2D LiDAR scans. We utilized ICP implementation from Open3D Python library. The results of volume measurements on the test dataset with ICP are MAE (MAPE): 0.95 m3 (9.41 %) and ME (MPE): – 0.46 m^3 (– 3.53 %) that are considerably high values. This might be the result of ICP sensitivity to outliers in the point clouds [20, 21], which is often the case in the considered dataset, containing plenty amount of challenging data. Also since we are dealing with moving objects, there is a lot of cluttering area in the LiDAR YZ scans during a truck pass. ICP is performing well in most passes, though in some cases absolute error surpasses 100 %, which results in high MAE on the dataset. Considering poor performance of ICP we utilized our own algorithm for displacement estimation on the LiDAR scans, showing better performance.

The utilized algorithm is based on consecutive LiDAR signals matching and is designed as follows. Two LIDAR YZ signals representing the truck's slices in the polar coordinate system are given and converted into a point cloud in the YZ plane. Signal matching is performed using edges on the observed surfaces. In order to extract data corresponding to the edges the points with derivatives in Z direction less than a threshold are removed. The optimal matching for the two 2D point clouds is estimated from which a displacement is obtained as a median distance between matched points.

4.6. The results

We have evaluated MAE and ME values for 5 volume estimation algorithm modifications: with LiDAR–based displacement estimator (LiDAR+LiDAR) and 4 vision–based with NeuFlow–v2, SIFT, ORB and ECO–TR. All competing algorithms were fine–tuned using the train dataset. The optimized loss function value (ME) for all competitors was less than 10^{-4} m^3.

Tab. 2 demonstrates the results of errors evaluation on the test dataset. As it can be seen from the Tab. 2 the top-3 algorithms ranking is the following: the LiDAR+camera with NeuFlow–v2 model is the most accurate demonstrating the lowest MAE value (0.57 m^3 on the entire dataset) and is followed by LiDAR+camera with ECO–TR (0.59 m^3) and LiDAR+LiDAR (0.63 m^3). The MAPE evaluations do not change the top–3 ranking. Overall from the experimental results it can be concluded that with MAE values the LiDAR+camera setup overcomes the LiDAR+LiDAR. Among vision–based competitors the best algorithm is NeuFlow–v2 estimating the dense optical flow that showed better results than the common sparse methods. It can be seen that accuracy of the LiDAR+camera with NeuFlow–v2 algorithm has a noticeable increase for passes of "dust" and "distorted" categories. Also it can be seen from Tab. 2 that LiDAR+LiDAR demonstrates smaller MAE values on the "dust" category than on the "clear" that can be explained with difference in data. If comparing all passes that were distorted with dust neglecting optical distortions (254 passes) with passes without LiDAR signal distortions (476 passes) than the difference in MAE becomes 0.85 m^3→ 0.53 m^3 (or 8.96 %→ 6.02 % for MAPE) correspondingly.

The ME values for passes from the "clear" category are comparable for the LiDAR+LiDAR and LiDAR+camera (NeuFlow–v2) configurations. The ME values for distorted data for the LiDAR+camera configuration are virtually identical to those for the undistorted, but for the LiDAR + LiDAR configuration, the value for the distorted is significantly lower. This may be the result of bias in the ground–truth from the true volume values. We plan to conduct studies with more precise ground–truth using calibrated objects of known geometry in future works.

Fig. 6 shows example significant LiDAR–based displacements estimation errors for a pass with sand dust following a passing truck. While the LiDAR–based algorithm detects displacements on sand dust, the NeuFlow–v2–based algorithm correctly determines the end of a pass which is indicated by rapid signal decrease (time interval from 2 to 3 seconds). By far no examples of extremely strong sand dust that may drastically decrease the visibility of the scene were met, nevertheless in strong sand dust cases the images quality may be enhanced if necessary. Since the model of image distorted by sand dust can be described by the hazed image model, restoration–based dehazing algorithms may be utilized [22, 23].

Tab. 2. MAE, ME, MAPE and MPE of cargo volume estimation evaluated on the large test dataset of 730 pairs of truck passes. L + L corresponds to LiDAR+LiDAR setup and L+C – LiDAR+camera

Algorithm	**All (730)**		**Dust (21)**		**Overexposure (163)**		**Shadow (17)**		**Clear (66)**		**Distorted (664)**	
	MAE, m^3 (MAPE, %)	ME, m^3 (MPE, %)	MAE, m^3 (MAPE, %)	ME, m^3 (MPE, %)	MAE, m^3 (MAPE, %)	ME, m^3 (MPE, %)	MAE, m^3 (MAPE, %)	ME, m^3 (MPE, %)	MAE, m^3 (MAPE, %)	ME, m^3 (MPE, %)	MAE, m^3 (MAPE, %)	ME, m^3 (MPE, %)
L+L	0.63 (7.00)	0.09 (2.57)	0.45 (5.55)	**–0.08** (1.04)	**0.55** (5.86)	0.23 (3.55)	0.48 (6.41)	0.25 (3.60)	0.57 (6.22)	0.31 (4.63)	0.64 (7.08)	0.07 (2.35)
L+C (NeuFlow–v2)	**0.57** (6.55)	0.32 (4.49)	0.33 (4.99)	0.12 (3.10)	0.57 (5.96)	0.32 (4.03)	0.53 (6.89)	0.51 (6.70)	**0.54** (5.90)	0.33 (4.59)	**0.57** (6.62)	0.32 (4.48)
L+C (SIFT)	0.74 (7.55)	–0.08 (–0.13)	0.48 (5.96)	–0.35 (–3.18)	0.87 (7.52)	–0.11 (–0.23)	**0.41** (4.68)	**–0.17** (–1.53)	0.67 (6.52)	0.06 (0.39)	0.75 (7.66)	–0.1 (–0.18)
L+C (ORB)	0.81 (8.20)	**–0.05** (–0.16)	0.55 (6.20)	–0.26 (–2.61)	0.89 (7.50)	**–0.02** (–0.10)	0.43 (4.89)	–0.22 (–2.16)	0.72 (7.27)	**–0.01** (–0.59)	0.82 (8.30)	**–0.06** (–0.11)
L+C (ECO–TR)	0.59 (6.75)	–0.32 (4.52)	**0.31** (4.37)	0.13 (2.85)	0.57 (6.07)	0.32 (4.04)	0.54 (7.11)	0.54 (7.03)	0.57 (6.22)	0.35 (4.82)	0.6 (6.81)	0.31 (4.49)

Among considered feature point–based algorithms the best results are shown by the ECO–TR model that are close in terms of MAE to that of NeuFlow–v2–based. On the one hand ECO–TR is a sparse alternative with similar performance to the dense NeuFlow–v2 algorithm but it has a sufficiently larger runtime. NeuFlow–v2 can be run in real–time on the

Quadro GV–100 GPU while the ECO–TR requires ~0.5 sec per frame on the same GPU. Training of light–weight descriptors is potentially a better alternative to computationally costly end–to–end models [24].

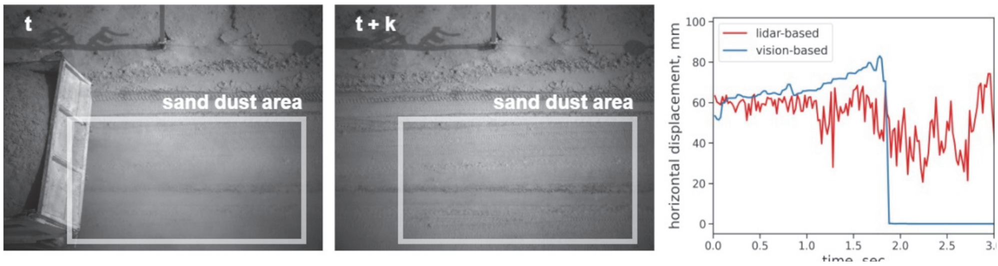

Fig. 6. Effect of inaccurate displacements estimation using the LiDAR YZ caused by sand dust presence. At the same time the sand dust is barely visible on images (sand dust area region) and the resulting estimates by the vision–based algorithm are much more accurate (the NeuFlow–v2–based algorithm was used)

Conclusion

The paper covers a problem of non–contact bulk cargo volume measurement for moving dump trucks in a semi–controlled outdoor environment and proposes a novel method that requires one single–layer (2D) LiDAR and a video camera. The displacements calculation method is based on dense optical flow estimation with NeuFlow–v2 model and conversion into metric shift value with help of the remaining LiDAR that provides depth information. Compared to the existing LiDAR+LiDAR state–of–the–art setup the proposed LiDAR+camera method is more accurate in terms of mean absolute volume estimation error. We have prepared a dataset containing 730 paired passes of various dump trucks on an operating sand pit annotated with a ground–truth volume value and demonstrate that the proposed LiDAR+camera method reduces the mean absolute error by 10 % (0.63 $m^3 \rightarrow 0.57$ m^3) where the baseline is the LiDAR+LiDAR setup. In addition to enhanced accuracy the proposed method utilizes a significantly cheaper sensor instead of a 2D LiDAR that noticeably reduces a total system cost. The future work is mainly focused on enhancing the fine–tuning and testing methodology using the calibrated cargo objects of known geometry, linear sizes and, thus, volume in order to reduce the impact of inaccuracies of bulk cargo volume estimation with existing techniques.

References

[1] Niskanen I, Duan G, Suzuki H, Endou D, Immonen M, Hiltunen M, Yamauchi G, Tyni P, Hashimoto T, Heikkilä R. Determining Payload on Platform of Lorry in Real Time Using Integrated 3–D Lidar From Excavator Boom. IEEE Transactions on Instrumentation and Measurement; 2024 Jan 5;73:1–7. DOI 10.1109/TIM.2024.3350119.

[2] Amorim LL, Mutz F, De Souza AF, Badue C, Oliveira–Santos T. Simple and Effective Load Volume Estimation in Moving Trucks using LiDARs. 2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Rio de Janeiro, Brazil: IEEE; 2019, p. 210–7. DOI: 10.1109/SIBGRAPI.2019.00036.

[3] Liu Y, Tao X, Li X, Colombo AW, Hu S. Artificial intelligence in smart logistics cyber–physical systems: State–of–the–arts and potential applications. IEEE Transactions on industrial cyber–physical systems. 2023 Jun 6;1:1–20. DOI: 10.1109/TICPS.2023.3283230. DOI: 10.1109/TICPS.2023.3283230.

[4] Sritap C, Pitayachaval P, Tantrairatn S. Development of a bulk material volume estimation system using automatic moving rail LiDAR technology. International Journal of Applied Methods in Electronics and Computers. 2024 Jun 30;12(2):48–53.

[5] Duff E, Automated volume estimation of haul–truck loads, Proceedings of the Australian Conference on Robotics and Automation. 2000, p. 179–184, CSIRO. DOI: 10.58190/ijamec.2024.97.

[6] Grigoryev A, Khanipov T, Koptelov I, Bocharov D, Postnikov V, Nikolaev D. Building a robust vehicle detection and classification module. InEighth International Conference on Machine Vision (ICMV 2015) 2015 Dec 8 (Vol. 9875, pp. 313–319). SPIE. DOI: 10.1117/12.2228806.

[7] Kokhan VV, Konyushenko ID, Bocharov DA, Seleznev IO, Nikolaev IP, Nikolaev DP. TSQ–2024: a categorized dataset of 2D LiDAR images of moving dump trucks in various environment conditions. InSeventeenth International Conference on Machine Vision (ICMV 2024) 2025 Feb 24 (Vol. 13517, pp. 60–65). SPIE. DOI: 10.1117/12.3055203.

[8] Zhang Z, Gupta A, Jiang H, Singh H. Neuflow v2: High–efficiency optical flow estimation on edge devices. arXiv preprint arXiv:2408.10161. 2024 Aug 19. DOI: 10.48550/arXiv.2408.10161.

[9] C. E. Sung, K. Young–Jin, J. Y. Joon, and J. S. Chan, Extraction of loaded cargo and volume calculation pcd for cargo vehicles. ICIC Express Letters Part B: Applications 2023 14(09), 977. DOI: 10.24507/icicelb.14.09.977.

[10] Savinetskij AB, Evstigneev VE, Kazaryan SM, Chudnikov VV, Bychkov AV, Rudzitis A, Kovalev AV. Method for determining the volume of bulk cargo in a moving vehicle using non–contact measurement. Russian Patent RU 2772138C1 of May 18, 2022.

[11] Liu Q, Feng C, Song Z, Louis J, Zhou J. Deep Learning Model Comparison for Vision–Based Classification of Full/Empty–Load Trucks in Earthmoving Operations. Applied Sciences 2019;9:4871. DOI: 10.3390/app922487.

[12] Sun X, Li X, Xiao D, Chen Y, Wang B. A method of mining truck loading volume detection based on deep learning and image recognition. Sensors. 2021 Jan 18;21(2):635. DOI: 10.3390/s21020635.

[13] Deniz M, Alam F, Yuan C, Ko HS, Lee HF. Single image based volume estimation for dump trucks in earthmoving using machine learning approach. EPiC Ser. Built Environ.. 2022 May;3:380–8.

[14] Kwak J, Bae J, Huh J, Moon DL, Hong D. Vision–based Payload Volume Estimation for Automatic Loading. Proc. of the 9th Intl. Conf. of Asian Society for Precision Engg. and Nanotechnology (ASPEN 2022) (2022). DOI: 10.3850/978–981–18–6021–8 OR–02–0280.

[15] Chen J, Lu W, Yuan L, Wu Y, Xue F. Estimating construction waste truck payload volume using monocular vision. Resources, Conservation and Recycling. 2022 Feb 1;177:106013. DOI: 10.1016/j.resconrec.2021.106013.
[16] Al–Qudah S, Yang M. Large displacement detection using improved lucas–kanade optical flow. Sensors. 2023 Mar 15;23(6):3152. DOI: 10.3390/s23063152.
[17] Lowe DG. Distinctive image features from scale–invariant keypoints. International journal of computer vision. 2004 Nov;60(2):91–110. DOI: 10.1023/B:VISI.0000029664.99615.94.
[18] Rublee E, Rabaud V, Konolige K, Bradski G. ORB: An efficient alternative to SIFT or SURF. In2011 International conference on computer vision 2011 Nov 6 (pp. 2564–2571). Ieee. DOI: 10.1109/ICCV.2011.6126544.
[19] Tan D, Liu JJ, Chen X, Chen C, Zhang R, Shen Y, Ding S, Ji R. Eco–tr: Efficient correspondences finding via coarse–to–fine refinement. InEuropean Conference on Computer Vision 2022 Oct 23 (pp. 317–334). Cham: Springer Nature Switzerland. DOI: 10.1007/978–3–031–20080–9_19.
[20] Deng Z, Yao Y, Deng B, Zhang J. A robust loss for point cloud registration. InProceedings of the IEEE/CVF international conference on computer vision 2021 (pp. 6138–6147).
[21] Brightman N, Fan L. A brief overview of the current state, challenging issues and future directions of point cloud registration. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2022 Oct 27;10:17–23. DOI: 10.5194/isprs–annals–X–3–W1–2022–17–2022
[22] Sidorchuk DS, Pavlova MA, Kushchev DO, Selyugin MA, Nikolaev IP, Bocharov DA. CADCP: a method for chromatic haze compensation on remotely sensed images. InSixteenth International Conference on Machine Vision (ICMV 2023) 2024 Apr 3 (Vol. 13072, pp. 346–355). SPIE. DOI: 10.1117/12.3023507.
[23] Filin AI, Kopylov AV, Gracheva IA. Method for removing haze from images, captured under a wide range of lighting conditions. Computer Optics. 2024;48(1):102–8. DOI: 10.18287/2412–6179–CO–1361.
[24] Volkov VV, Shvets EA. Neural network algorithm for optical–SAR image registration based on a uniform grid of points. Computer Optics. 2024;48(4):610–8. DOI: 10.18287/2412–6179–CO–1426.

About authors

Dmitry Aleksandrovich Bocharov (b. 1992), received the Ph. D. degree in computer science, in 2022. He is currently a Researcher with the Vision Systems Laboratories in Institute for Information Transmission Problems and Federal Research Center ‘‘Computer Science and Control’’ of the Russian Academy of Sciences, with research interests focused on computer vision and industrial recognition systems. E–mail: *bocharov.mitry@gmail.com*

Vladislav Vladimirovich Kokhan (b. 1995), graduated from Moscow Aviation University in 2019. Works as a researcher at the Vision Systems Laboratory of IITP RAS. His major research interests include image processing and computer vision. E–mail: *vkokhan1@gmail.com*

Ivan Denisovich Konyushenko (b. 2003), graduated from Moscow Institute of Physics and Technology, in 2025. He is currently a research intern at the Vision Systems Laboratories in the Institute for Information Transmission Problems of the Russian Academy of Sciences, with research interests focused on deep learning and computer vision.
E–mail: *idkonyushenko@gmail.com*

Artem Yurievich Resniansky (b. 1972), graduated from Moscow State University department of Physics in 1994. Works as a software engineer in the Federal Research Center “Computer Science and Control’’ of the Russian Academy of Sciences. His research focuses on design, development, and application of systems for acquiring, processing, and interpreting digital data related to optical measurements. E–mail: *artem.res@gmail.com*

Ilya Petrovich Nikolaev (b. 1972), Ph. D. in Physics and Mathematics. He graduated from Lomonosov Moscow State University in 1994, majoring in Physics. After years of working in the field of adaptive optics, in 2020 he joined the vision systems laboratory at the Institute for Information Transmission Problems (IITP RAS). The research interests are image processing and visual perception. E–mail: *i.p.nikolaev@iitp.ru*

Dmitry Petrovich Nikolaev (b. 1978), received the Doctor of Sciences degree in computer science, in 2023. Since 2007, he has been the Head of the Vision Systems Laboratory, Institute for Information Transmission Problems of the Russian Academy of Sciences, Moscow. He has been the CTO of Smart Engines Service LLC, Moscow, since 2016. Since 2016, he has been an Associate Professor with Moscow Institute of Physics and Technology, State University, Moscow, teaching the image processing and analysis course. Currently, he is the Head of the Vision Systems Laboratory, Federal Research Center ‘‘Computer Science and Control’’ of the Russian Academy of Sciences. He has authored more than 190 Scopus–indexed publications and 15 patents. His research interests include computer vision and color image understanding. E–mail: *d.p.nikolaev@gmail.com*

Received August 12, 2025. The final version – September 06, 2025.