Computer Vision and Deep Learning for Real-Time Pavement Distress Detection

72

Kristina Doycheva, Christian Koch, and Markus König

Abstract

Despite of the increased level of automation in vehicles, the detection of pavement distress, such as cracks and potholes, is mostly performed manually. We propose a methodology for automated pavement distress detection based on computer vision. Thereby, images obtained by cameras installed in common passenger vehicles are analyzed in real time, resulting in cost savings and a reduced amount of stored data. For this purpose, the wavelet transform was implemented on Graphics Processing Units (GPU). In addition, median filtering and top-hat transform were also implemented on GPU to enable real-time noise removal and correction of non-uniform background illumination. To distinguish between surface types, we incorporated textural features into our methodology and deep learning was utilized to determine the distress type (cracks, potholes or patches). Results obtained by different vehicles were aggregated to improve the reliability of the methodology. Case studies were conducted for validation and tests achieved promising results.

Keywords

Pavement distress detection • Graphics processing units • Textural features • Deep learning

72.1 Introduction

The number of cars on the road is expected to double by 2040 [1]. This expectation is based on the fact that the number of passenger vehicles registered in 2017 increased compared to 2016 [2]. For example, the total number of passenger vehicles registered in Germany on January 1, 2017 was 45,803,560 (1.6% more compared to 2016). As passenger cars fulfill the desire for mobility and flexibility, they accounted for 83.4% of inland passenger transport in the European Union in 2014 [3].

In order that the reliability of cars as a means of passenger transport remains high, infrastructure in good condition is essential. However, in recent years, the condition of municipal roads has deteriorated. As a result, pavement distress, such as cracks, potholes and patches is visible on the road surface. This distress affects the driving quality and leads to traffic congestions. Furthermore, vehicle damage due to accidents caused by distress such as potholes is often incurred.

To extend the life of the road pavement, maintenance activities need to be planned carefully and executed timely. Thereby, maintenance activities should be frequent and cost-effective [4]. Surface treatments, such as pothole repair and crack sealing, are amongst the most important pavement maintenance activities. To allow surface treatment, the precise location of distress should be known. To this end, the condition of the road surface should be frequently evaluated. The presence of defects, such as cracks, potholes and patches, should be detected and the extent and severity of the distress should be estimated to allow planning repair actions appropriately.

K. Doycheva (✉) · M. König
Ruhr-University Bochum, Universitätsstr. 150, Gebäude IC 6-71, 44801 Bochum, Germany
e-mail: kristina.doycheva@rub.de

C. Koch
Bauhaus-University Weimar, Marienstr. 13a, Room 208, 99423 Weimar, Germany

I. Mutis and T. Hartmann (eds.), *Advances in Informatics and Computing in Civil and Construction Engineering*,
https://doi.org/10.1007/978-3-030-00220-6_72

72.2 Related Work

Nowadays, the detection of distress is mostly performed manually by teams of inspectors who observe the condition of the roads while driving. However, this method is costly if we consider the extent of the existing road network. Moreover, it is also subjective, because it is based on the evaluation by a single expert. A further critical drawback of the manual evaluation is the need to provide a safe surveying environment by controlling the traffic during an inspection.

Due to these and other drawbacks, an automated pavement distress detection approach is necessary.

Several publications propose such approaches. Some of them are based on sensors such as accelerometers, but the majority utilizes computer vision to detect distress in pavement images. The vision-based approaches are usually based on the assumption that pixels, which belong to the distress area, have different intensities compared to the pixels in the background (i.e., the intact pavement) [5].

While only a few methods are capable of detecting different types of distress [6, 7], most approaches are specifically developed for a particular type of distress. Commonly, statistical properties of the images are used to derive features for distress detection. Machine learning algorithms employ these features to build classifiers that can distinguish between images of intact pavement and images containing distress. Approaches for crack detection, which are based on machine learning, have been presented [8–10]. Another approach was proposed by Zou et al. [11], who analyze the difference in the intensities of regions in an image and derives a crack probability map based on tensor voting. Approaches for potholes have also been proposed [12, 13]. A small number of methods for patch detection have been developed. Cafiso et al. [14] have proposed a method for analysis of images with respect to patches based on clustering. Radopoulou et al. [15] have applied morphological operations to segment patch regions. All publications cited above report promising achievements and classification results. Nevertheless, they suffer from some issues and limitations. For example, a huge amount of data must usually be stored persistently and processed offline. In addition, collecting data for sensor-based methods, such as methods based on accelerometers, affects passengers comfort while driving, because it is required to drive straight over the distress.

Collecting data for computer-vision based methods, in contrast, can be per-formed while avoiding driving over the distress. However, methods based on computer vision strongly depend on the weather condition and daytime the images have been obtained at, because they could lead to non-uniform illumination. Moreover, most methods cannot be applied in general, because they do not distinguish between different pavement types and are optimized for a specific pavement surface.

72.3 Research Questions and Objectives

This work aims to compensate these deficiencies by answering the following research questions:

- How can pavement distress be both reliably and cost-effectively detected without affecting traffic?
- In case of automated distress detection, how can the amount of stored data be reduced?
- How can distress be detected on various types of pavement surfaces under diverse lighting conditions?
- Is it possible to distinguish between different distress types and, if so, how can that be achieved?
- How can distress locations be determined precisely using inexpensive hardware?

To answer these research questions, the objective of this work is to propose a cost and storage efficient automated approach for pavement distress detection. This approach should minimize classification subjectivity and allow driving with usual speed with minimal driving restrictions. Moreover, distress should be detected under various lighting conditions and on different types of roads, whereby the type of distress should also be determined. In addition, the distress location and the severity or extent of the distress should be estimated and results obtained from different vehicles should be aggregated in order to guarantee reliability.

72.4 Methodology

In order to answer these research questions and to achieve the objective presented above, in this paper the following concept is proposed (Fig. 72.1).

Since nowadays there is a trend that new vehicles are equipped with rear view cameras [16], common vehicles such as taxis and buses equipped with such cameras and Global Positioning System (GPS) receivers can be used instead of inspection vehicles. The cameras obtain images of the road surface while the vehicles are driving at their usual speed on their usual route. Although the quality of the images may be worse than the quality of images taken by special inspection vehicles, images taken by various vehicles at the same location are used in order to increase the reliability of the detection.

The distress detection is distributed and based on two stages, namely rough detection and ne detection. In the rough detection stage, vehicles constantly obtain images of the pavement surface of parts of the road network. These images are analyzed in real time with respect to pavement distress. If distress is detected on an image, the classification label of the distress type and the GPS coordinates of the vehicle at the time the image was taken are saved persistently and forwarded to a central server. The potential distress locations identified by different vehicles are clustered and requests containing their GPS coordinates are sent to the ne detection vehicles in the second stage.

The fine detection vehicles activate their cameras only when they approach potential distress locations. Algorithms that are more complicated than the ones utilized in the rough detection stage are then applied on the images in order to determine automatically the severity and extent of the distress. Thereby, different algorithms are used by different vehicles, such as taxis and buses, whereby each vehicle receives a request form the server containing a list of potential distress locations and the id of the algorithm that has to be used by that vehicle. At the end, the results are combined, so that the final evaluation is based on numerous results and should, thus, be accurate. Moreover, the results of the ne detection stage are used to improve the rough

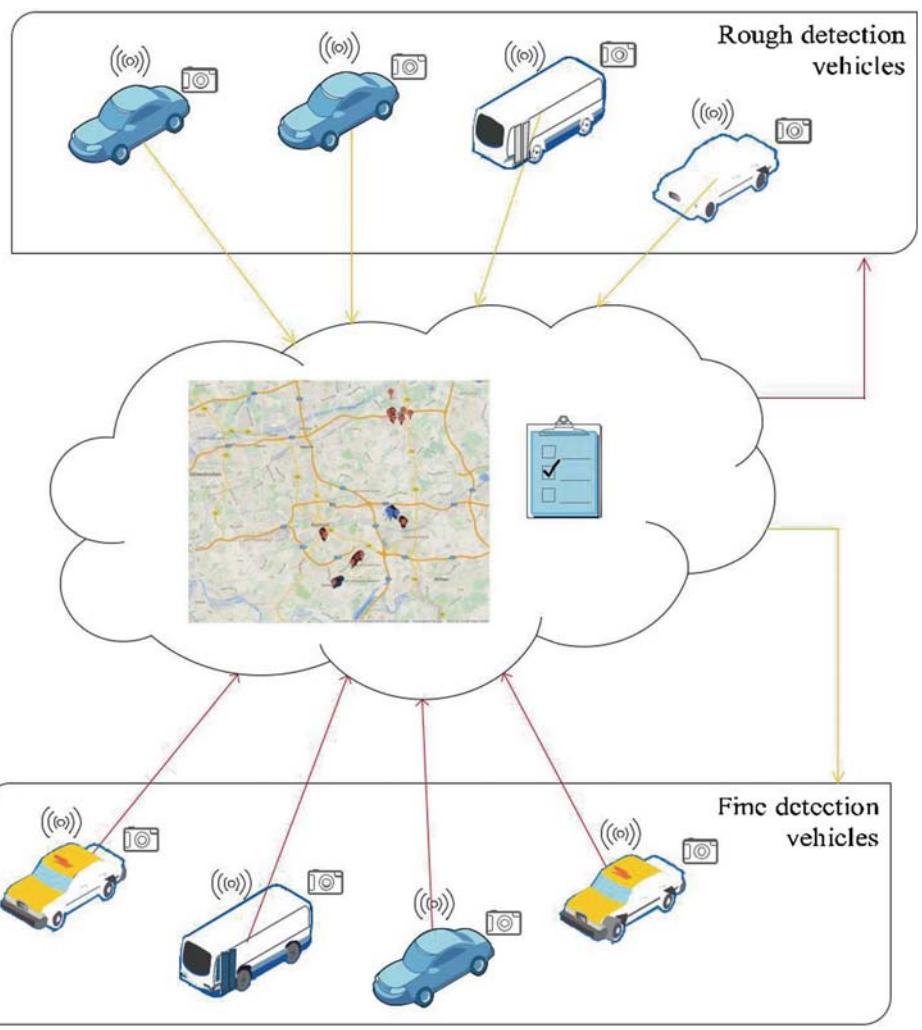

Fig. 72.1 Distress detection concept based on common vehicles

detection. For example, if a crack image was misclassified in the rough detection stage, the correct classification label of the distress at the specific location by the ne detection vehicle is used to update the classification model.

This paper focuses mainly on the rough detection stage, which has been fully implemented. An approach for clustering locations determined by different vehicles has also been developed and implemented. Several state-of-the-art algorithms for fine detection have been adopted and implemented, but they have not been extensively tested yet.

72.5 Implementation

72.5.1 Rough Detection

An overview of the rough detection approach is presented in Fig. 72.2. The rough detection method is based on the wavelet transform and it was implemented on a GPU in order to allow real-time distress detection, as described in [17]. The wavelet transform has been used by Zhou et al. for pavement distress image classification [7]. Nevertheless, if it is executed on a CPU and implemented using sequential code, the wavelet transform and the calculation of features for classification based on it require too much time and make it inpossible to perform image analysis in real time. Prior to applying the wavelet transform on the images, the latter are pre-processed to remove noise and shadows. For this purpose, median filters and top-hat transforms were also implemented on the GPUs.

Although the wavelet analiysis leads to good classification results, it is not capable of detecting distress on various types of road surfaces, because the values of the classification features strongly depend on the surface texture. To incorporate pavement surface texture characteristics, textural features were integrated. The calculation of the textural feature values was carried out on GPUs using the Open Computing Language (OpenCL), as presented in [18]. In particular, four features proposed by Haralick [19] and computed based on the gray-level co-occurence matrix (GLCM) of an image are calculated and used to enhance the distress detection.

Based on the wavelet transform and the textural features, distress can be detected reliably in images of various types of road surfaces. However, to accurately evaluate the state of the pavement surface, it is necessary to also determine the type of the distress, i.e., to distinguish between cracks, potholes and patches. To this aim, state-of-the-art concepts from deep learning were incorporated. Specifically, a classification model was generated using AlexNet, a deep convolutional neural network developed by a team at the University of Toronto [20]. In addition to determining the category an image most probably belongs to, the network associates a probability value to the prediction. In our methodology, this probability is taken into account by being compared to a value above which the result is assumed to be reliable. This threshold value is derived for each class based on experimental results, as described in the following section.

Finally, if sufficient training data is available for the specific conditions and the wavelet-based method can be applied, the results of the methods described above are combined. As a consequence, the confidence that the final rough detection result is correct increases greatly.

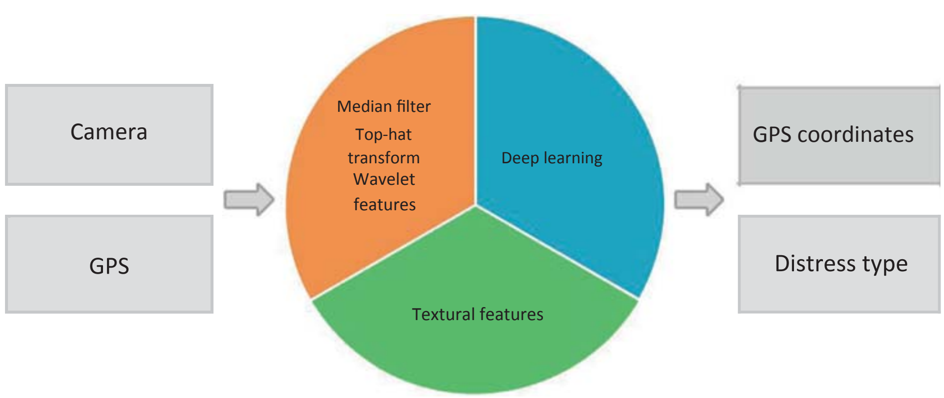

Fig. 72.2 Rough distress detection approach

72.5.2 Fine Detection

The fine detection is based on a suite of algorithms specifically developed for different types of distress. Since the algorithms perform differently well depending on the lighting condition, extent of the distress and the available training data, the ne detection is based on ensemble learning. The idea behind ensemble learning is that by combining multiple classification models, a model which out-performs the individual models is obtained [21]. The models can be combined by using methods such as majority voting, weighted averaging, stacking, etc. Different analysis algorithms will be investigated and implemented in the future. In addition, various ways to build an ensemble model based on these algorithms will be examined and evaluated.

Moreover, with the recent advances in deep learning [22], convolutional net-works are capable of not only classifying an image into one of many possible categories, but also detecting (localizing) a single object within an image and determining its size as well. Future work includes examining the possibility of using deep learning to estimate the severity and extent of the distress.

72.5.3 Georeferencing and Location Clustering

To allow the subsequent maintenance of roads, it is required that the locations of all detected distresses are known. To this end, GPS receivers capable of determining the geographical coordinates (latitude and longitude) of a vehicle are installed. Using these coordinates, the images of a potential distress can be georeferenced. To increase the reliability of the detection, repeated measurements or analyses can be performed by various vehicles using different algorithms. The results of these analyses are then combined to estimate the location of the distress more precisely. Thus, errors due to GPS inaccuracy are avoided. Moreover, outliers, i.e., single non-repeated and thus probably incorrect results, are eliminated.

In this work, the GPS coordinates are processed using a clustering method. The DBSCAN algorithm [23] was applied to identify the clusters, assign the distress coordinates to these clusters and eliminate outliers. Afterwards, the precise distress location is determined using statistical evaluation. The average values of the latitude and longitude coordinates are calculated from all points belonging to the same cluster. Finally, statistical evaluations and the number of detected distress instances per road section are used to determine which sections of the road are in poorest condition.

72.6 Case Studies

72.6.1 Validation

To validate the approach presented in this thesis, several case studies were con-ducted. A high-frequency camera was installed on a rear-door back-carrier of a car as shown in Fig. 72.3. In addition, a GPS receiver was mounted on the top of the car. Using the camera, more than 50,000 images of the road surface in Bochum and Witten, Germany were obtained. However, some of these images contained more than one type of distress. In total, 37,982 images contained either cracks (6419 images), patches (5437 images), potholes (150 images) or no distress (12,682 images). These images were used to test the rough detection methodology. The images were randomly split, so that 90% were used for training and the remaining 10% were used for validation. A classification accuracy of 88% was achieved. However, the training dataset contained a small number of pothole images and as a result the pothole images were mostly classified incorrectly. The accuracy can be improved by using a more appropriate dataset.

Furthermore, threshold values above which the classification is assumed as reliable were calculated. For this purpose, the average probability values for the correctly and incorrectly classified images for each class were computed. For example, in case of cracks the average confidence for incorrectly predicted classes was equal to 70%, while the probability associated to correct predictions was significantly higher (85%). Due to this reason, it could be assumed that images classified as containing cracks with a probability higher than 77% presumably do contain cracks.

The location clustering was tested using a small case study. A vehicle obtained images of a road section in Bochum, Germany, twice and then analyzed them with two different algorithms. In total, 642 distress locations were detected. Out of these, 382 locations were detected by the first algorithm and the remain-ing 260 locations were identified by the second algorithm. After applying the DBSCAN algorithm, 122 clusters were generated. The clusters were built under the condition that a cluster must contain at least two points, whereby the maximum distance between two points is 5 m. The results indicate that distress was incorrectly detected by the first algorithm at locations where distress was not actually present.

Fig. 72.3 Vehicle used for the case studies

These points were determined by the DBSCAN algorithm as outliers. In the next step, the distress location was estimated by means of statistical evaluation.

72.6.2 Performance Evaluation

Performance tests were carried out to evaluate the speed-up achieved by utilizing GPUs for the implementation. Images with di erent sizes, namely 256 $\times$ 256, 512 $\times$ 512, 1,024 $\times$ 1,024, and 2,048 $\times$ 2,048 pixels, were used. A 2.10 GHz Intel Core i7-4600 CPU was used to execute a sequential version of the code and the parallel version was executed on an Nvidia Tesla C2070 GPU. The highest speed-up was achieved for images with 2,048 $\times$ 2,048 pixels. In case of the median filter, top-hat transform and wavelet transform, the overall speed-up was 9,009, while the GPU execution of the calculation of the gray-level co-ocurrence matrix and the textural features was 81 times (for the GLCM), respectively 1381 times (for the features), faster than the sequential execution.

72.7 Summary and Outlook

In this thesis, a methodology towards pavement distress detection has been presented. The methodology is based on computer vision. In particular, the wavelet transform and textural features of images obtained by a camera installed in a vehicle were used to detect potential distress such as cracks, potholes and patches. The analysis was implemented on GPU to allow real-time distress detection while the vehicles are driving. In addition, a GPU implementation of the median filter was adopted and top-hat transform was implemented on GPU to pre-process the images by removing noise and correcting non-uniform background illumination.

To determine the distress type, deep learning was utilized. As a result, it is possible to classify images into one of the four categories intact pavement, crack, pothole, and patch. The images were georeferenced using a GPS receiver installed on the top of the vehicle and a precise GPS location was estimated by clustering potential distress locations detected by different vehicles.

The methodology is based on the idea that common passenger vehicles are used instead of special surveying vehicles. Although the quality of the data obtained by passenger vehicles may not be as high as the quality of the data obtained by special vehicles, reliable detection of the distress is possible, because more than one vehicle executes the same analysis algorithm for the same road. In addition, results produced by different methods are combined.

The case studies have proven that by exploiting GPUs, pavement distress detection can be performed in real time. The integration of the textural features allowed detecting distress on various types of road services. It has been demonstrated that deep learning accurately classifies images according to the distress type, which is important in order to plan further maintenance actions properly.

Recent developments also allow further refinements of the methodology. For instance, deep learning is not only capable of classifying an image into different categories, but of localizing objects within an image. The latter could be used in the future to determine the extent of the distress. Also, the information provided by automated pavement distress detection can be used not only by municipalities in order to schedule further maintenance actions, but also by ordinary citizen by warning them when they approach severe distress that has not been fixed, for example. Future research may also include integrating pavement distress data into Building Information Models (BIM) for roads, which can be used to support management processes.

References

1. Smith, M.N.: The number of cars worldwide is set to double by 2040, April 2016. https://www.weforum.org/agenda/2016/04/the-number-of-cars-worldwide-is-set-to-double-by-2040
2. Bekker, H.: Germany: Total number of registered cars, March 2017. https://www.best-selling-cars.com/germany/2017-germany-total-number-registered-cars/
3. Eurostat: Passenger transport statistics, April 2018. http://ec.europa.eu/eurostat/statistics-explained/index.php/Passenger_transport_statistics
4. Pavement Interactive. Pavement maintenance—prevention or repair?, May 2013. http://www.pavementinteractive.org/2013/05/27/pavement-maintenance-prevention-or-repair/
5. Koch, C., Georgieva, K., Kasireddy, V., Akinci, B., Fieguth, P.: A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure. Adv. Eng. Inf. **29**, 196–210 (2015)
6. Lokeshwor, H., Das, L.K., Goel, S.: Robust method for automated segmentation of frames with/without distress from road surface video clips. J. Transp. Eng. **140**(1), 31–41 (2014)
7. Zhou, J., Huang, P.S., Chiang, F.-P.: Wavelet-based pavement distress detection and evaluation. Opt. Eng. **45**(2), 027007 (2006)
8. Li, L., Sun, L., Ning, G., Tan, S.: Automatic pavement crack recognition based on BP neural network. Promet Traffic Transp. **26**, 11–22 (2014)
9. Moussa, G., Hussain, K.: A new technique for automatic detection and parameters estimation of pavement crack. In: 4th International Multi-Conference on Engineering Technology Innovation (IMETI 2011), 2011
10. Varadharajan, S., Jose, S., Sharma, K., Wander, L., Mertz, C.: Vision for road inspection. In: Proceedings of WACV 2014: IEEE Winter Conference on Applications of Computer Vision, 2014
11. Zou, Q., Cao, Y., Li, Q., Mao, Q., Wang, S.: Cracktree: automatic crack detection from pavement images. Pattern Recog. Lett. **33**, 227–238 (2012)
12. Koch, C., Brilakis, I.K.: Pothole detection in asphalt pavement images. Adv. Eng. Inf. **25**, 507–515 (2011)
13. Li, Q., Yao, M., Yao, X., Xu, B.: A real-time 3d scanning system for pavement distortion inspection. Meas. Sci. Technol. **21**(8), 015702 (2010)
14. Cafiso, S., Di Graziano, A., Battiato, S.: Evaluation of pavement surface distress using digital image collection and analysis. In: Seventh International Congress on Advances in Civil Engineering, 2006
15. Radopoulou, S.C., Brilakis, I.: Improving patch distress detection using vision tracking on video data. In Proceedings of the 21st International Workshop on Intelligent Computing in Engineering, 2014
16. National Highway Traffic Safety Administration (NHTSA). Federal motor vehicle safety standards: rear visibility, 2014
17. Doycheva, K., Koch, C., König, M.: GPU-enabled pavement distress image classification in real time. J. Comput. Civ. Eng. **31**(3) (2017)
18. Doycheva, K., Koch, C., König, M.: Implementing textural features on gpus for improved real-time pavement distress detection. J. Real-time Image Proc. (2016)
19. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **3**(6), 610–621 (1973)
20. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classi_cation with deep convolutional neural networks. In NIPS 2012: Neural Information Processing Systems, Lake Tahoe, Nevada (2012)
21. Rokach, L.: Ensemble-based classifiers. Artif. Intell. Rev. **33**(1–2), 1–39 (2010)
22. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**, 436–444 (2015)
23. Ester, M., Kriegel, H.-P., Sander, J., Xu, X.I.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: Simoudis, E., Han, J., Fayyad, U.M. (eds.) Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96) (1996)