Pre-Study on Automatically Determining Road Condition with a Camera

Cris Luengo
Centre for Image Analysis, Uppsala University
Box 337, 751 05 Uppsala, Sweden
cris@cb.uu.se

5 July 2013

Contents

1 Introduction

The study reported herein was commissioned by Jonas Hallenberg and Pertti Kuusisto at Trafikverket, with the aim of discovering what is the state of the art in automatically determining the road condition (or road weather) from an image recorded by the new camera system (Sony SNC-CH180) that is being installed along roads all over Sweden. Currently, images are sent to a central location, and displayed on computer monitors to operators. Given that in February 2013 there were 250 cameras operational, with plans to add another 300 cameras, it is only logical that automated analysis of the images is in order.

Of interest is, foremost, to detect if the road is covered with snow or not. But also important would be to distinguish dry, wet, and icy road conditions. During the literature study I've found group in Canada that works on distinguishing five road conditions: bare dry, bare wet, thin snow cover, slushy snow cover, partially snow cover, mostly snow cover [5]. They try to distinguish these conditions with friction measurements, but recorded video of the road at the same time, and had manual operators mark each section of road measured to be in one of the five categories. They found that it is difficult to visually distinguish thin snow cover from slushy snow cover, and slushy snow cover from partially snow cover, due to slushy snow cover being somewhere in between the other two types in appearance. This leads me to assume that, if this many categories would be required, a system using only a camera will not produce very good results.

Part of the requirements for this study was to interview Jörgen Bogren (Department of Earth Sciences, University of Gothenburg) and Patrik Jonsson (Mid Sweden University, Östersund, and Combitech AB). I had a long and interesting discussion with Patrik Jonsson, that helped shape the recommendations given in this report (Section 3). However, I have not been able to contact Jörgen Bogren, either through email or phone. I did, however, read about his research regarding modelling of local variations in surface temperature, which could be useful when deciding where to place the cameras, so as to look at the part of the road most likely to be slippery, but will not be as useful as a simple thermometer when it comes to obtaining information to aid in analysing images (more on this in Section 3).

2 Current State of the Art

Neither I nor Patrik Jonsson have been able to identify a commercial system capable of recognising the road condition through a photographic or video camera, although there are systems that use other type of sensors, and that are good enough to be used as "the truth" in at least one study: Jokela et al. [8]. For example, Vaisala Oyj produces a "remote surface state sensor" (DSC111), which measures the backscattered signal of an infrared light beam [4].

There is very little in the scientific literature regarding the recognition of road condition through a photographic or video camera. The first part of this section summarises the studies I found. Interesting to note is that most of the research in this area is performed in Sweden and Japan. One group in Finland has also produced some papers, and I found one paper each from Canada and China.

In this section I use the terms *classification* and *feature*. All of the methods described below classify images into one of several classes, which means that, for each image, they decide whether the road has water, ice, snow, etc. In this case, the classes are dry, wet, icy, snowy, etc. A classifier is a program that uses features as input, and a class label as output. Features are quantitative data derived from the image, such as the intensity at a certain point in the image, the difference in intensity between two points, etc. A classifier is trained by using a set of images with known classes. When testing the classifier to see how well it works, it is important to use new images, that have not been used in the training. This is because otherwise it is not possible to know how well the classifier is able to generalise from the examples. It is very possible for a classifier to only produce good results for those example images in the training set, performing very poorly on other images. Finding the right features to put into the classifier is of course very important, but so is finding the right classifier that works well with the structure of the given features.

2.1 The Studies

The oldest study I found as that of Kuehnle and Burghout [15], at Dalarna University, which, using a hand-held camera next to a road, some rather simple image analysis techniques, and artificial neural networks (ANN) obtained an overall classification accuracy of 40–50%, not being able to distinguish ice from wet road or tracks. McFall continued this study, showing that a similar system using a microphone as input yields 88% classification accuracy [18], and combining the two systems yields a near perfect classification [19]. A study from Japan found a 90% classification accuracy with a road-side microphone [14].

At Mid Sweden University, Jonsson [9, 10] also combined camera images with other information, yielding results much better to those of camera information alone. Features like road surface temperature and dew point add significantly to the ability to predict the road condition. The image analysis methods that Jonsson used for feature extraction is from a paper published in 1991. More advanced feature extraction methods should improve the classification significantly. He also suggests the use of an infrared detector to further improve the results [11]. Infrared frequencies have long been known to be useful in distinguishing ice, water and snow.

Zhang et al. [27] use features derived from the co-occurrence matrix, also a very old technique, and reach a classification accuracy of 93%, but do not try to distinguish wet or icy roads, only heavy snow cover, mild snow cover, and dry. Omer and Fu [21] use different features, but also quite simple ones, with images of various vehicle-mounted cameras, yielding a classification accuracy of 81–89% for distinguishing what they call bare, covered and tracks.

None of the studies discussed so far take into account the illumination: the same scene can look very different to a camera if it is sunny, cloudy, or night time. Riehm [22], who uses average image intensity to estimate the amount of snow on the road, quickly comes to this conclusion as well. One study by Takeuchi et al. [24] uses images taken at daytime and nighttime, training separate classifiers for day and night, as well as a single classifier that is capable of classifying images irrespective of the time of day. They report a classification accuracy of 66% at night and 85% at day (similar for both approaches) for distinguishing wet road from snow-covered road. However, it is not clear from the paper how they divide their rather small dataset into a training and a testing set. The nighttime images are illuminated by standard road-side illumination. Interesting to note is that the white line in the middle of the road tends to be classified as snow covered, showing the importance of having a correct model: if one assumes that the bare road is black, the road markings do not fit the model.

From the same group is a paper [7] with a similar method, but interested only in nighttime road conditions. In this paper they use the illumination from car headlights as they drive past. Results on several video frames need to be combined to obtain information from the whole road, as the headlights illuminate only a relatively small portion. They obtain recognition rate of less than 75% for distinguishing dry and wet road. A third paper by the same group [12] reports the same nighttime-only methodology, but with a vehicle-mounted camera, in which the illumination is much less consistent, nonetheless reporting an accuracy of 90–96% for distinguishing dry, wet and snow (the evaluation methodology is also not clarified in this paper, and I'm certain something is not correct, as these results do not match those of the other two papers from the same group).

If one is not limited to a single camera, it is possible to compare the radiance of a scene under two orthogonally oriented polarising filters. It is well known that light specularly reflected off a surface is polarised according to the orientation of that surface. One of the two cameras would see the specular reflection off the wet road, whereas the other one would not. This seems to have first been explored to detect wet road surface by Ueda et al. in 1994 (I have not been able to find this paper except in references from papers by the same author). Yamada et al. [25] combine this method and features extracted from the co-occurrence matrix to get an average classification accuracy of 92%, distinguishing dry, wet, slushy, icy and snowy pavement. The same group also published these methods applied to a vehicle-mounted camera [26]. Kutila et al. [16, 17] proposed a very similar system, but using a different texture measure, and reported a 93% accuracy in detecting ice, but only a 61% accuracy in detecting water. This group also has a vehicle-mounted variant [8].

2.2 Relevant Methodology

2.2.1 Classifiers

Many of the papers above use either artificial neural networks (ANN) or support vector machines (SVM) as a classifier. ANNs have been around since before the digital era, when they were trained by manually turning nobs on variable resistors. They have been in and out of fashion multiple times, and currently seem to be on a comeback. However, still little is understood regarding how to best select the network configuration. A typical ANN has a set of input nodes, one or more "hidden layers" (a set of internal nodes), and a set of output nodes, typically one for each class. Each node computes a weighted average of all their inputs, applies a nonlinear function to this average, and the result is presented at the output. Each node in the input layer has all features as inputs, and each node in the other layers is connected to all nodes in the previous layer. The result is a highly non-linear operation with a huge amount of parameters to set during training, and a classifier with a very complex decision boundary. Many training methods exist, all of which iteratively tweak the parameters to improve the classification. The more modern ones take care of not over training, that is, making sure that the ANN is still generalising from the input. Because of the huge amount of parameters, over training can happen easily.

SVMs are currently considered the state-of-the-art classifier. These can be trained by a simple minimisation operation, and therefore are much easier and quicker to train than ANNs. In that sense, they behave much more like the traditional classifiers (Naive Bayes, Fisher's, quadratic, etc.). However, as opposed to these traditional classifiers, SVMs do not try to to estimate the distribution of the data (such as fitting a Gaussian distribution to the data), and instead just find the linear classification boundary that maximises the distance to all samples and completely separates the samples of the two classes. Note that, if more than two classes exist, multiple SVMs need to be trained to distinguish all the classes. Non-linear SVMs can be generated through the kernel technique, where non-linear combinations of features are added to the feature set. Because the method does not try to estimate the distribution of the data, it is the only method that seems to work well if there are relatively few training data with respect to the number of features.

Another popular classifier is AdaBoost [6]. This classifier combines many very simple two-class classifiers in a chain, such that later classifiers only need to classify the part of the input that all previous links classified as class B. The first classifier will classify a small portion of the input as class A, with very high certainty. The next classifier no longer needs to look at these inputs, so can use a different combination of features to again find a few inputs to classify as class A with very high certainty. The more stages the chain has, the more complex the decision boundary becomes. The difference with ANN is that AdaBoost is less likely to become over trained, and the number of parameters depends on the complexity of the data, rather than the a priory choice on network configuration.

For a very complete overview of these techniques see Bishop [2].

2.2.2 Feature Extraction

In the studies discussed above, only rather simple image analysis methods have been used for feature extraction. A co-occurrence matrix is a table in which the element at position (i, j) counts how many times in the image there is a transition from intensity i to intensity j. Thus, this matrix contains information comparing every pair of pixels situated next to each other. The matrix for horizontal transitions will be different from that for vertical transitions, and it is possible to create such a matrix also for pixels with any other relative location, for example where the second pixel is three steps right and two down from the first one. The longer the distance between the two pixels, the less information the matrix contains, as correlation typically goes down significantly. However, for all four neighbouring positions, such a co-occurrence matrix has been shown to give information regarding the texture in the image (note that, of all 8 primary directions in the square grid formed by the pixels, four directions give the transposed of the matrix given by the other four directions, and therefore add no new information). However, it is statistics derived from the matrix that are used as texture features, never the matrix itself.

Other features used, such as the average intensity, gradient strengths, etc. depend very much on the illumination in the scene, and cannot be considered suitable features for the task at hand. The advantage of texture features is that they can be made more or less independent of the illumination intensity. The texture can still change, however, with the illumination direction, if the texture seen in the image is the result of shadows cast by a rough surface. In such a case, on a cloudy day there will be no or little shadow, and thus less texture in the image compared to a sunny day. Likewise, the texture will look very different if illuminated by the Sun in the morning or the afternoon, or by the camera's NIR illumination at night.

More recent, and highly successful, texture measures are local binary patterns [20] and all the derived methods. For these methods, a number of pixels, usually six or eight, in a circle around a central pixel are examined. If the value of a pixel is higher than that of the central pixel, it is given a value of one, otherwise it is given a value of 0. All six or eight digits are then combined into a single binary number. This is done for all pixels in the image. A histogram of the resulting values is then used as a feature vector describing the texture in the image. Some simplifications are often applied to reduce the number of features produced. To make the method rotationally invariant (i.e. to make the method produce the same features for a rotated image), the binary digits can be circularly shifted to be as low as possible. Another method circularly shifts the histogram instead. These two approaches to rotational invariance are useful under different circumstances.

2.2.3 Pre-Processing

Rather than throwing a feature extraction algorithm at the raw image data, it makes sense to pre-process the input image first, extracting only the relevant portions that show the road and discarding everything else. Such preprocessing can also detect whether the image has a clear view of the road, or if, for example, a blizzard is obscuring the view.

The first step here would be to manually create a mask image (this can be done at the time that a camera is first installed) that shows which areas of the image display the road. To exclude areas of the road covered by cars, a background estimation technique like Mixture of Gaussians (MoG) can be used [3, 23]. This technique works by comparing images over a long period of time, and decides what the scene looks like without any moving objects in it (usually referred to as the background). MoG adapts its concept of the background through changes of illumination, which typically are much slower than the cars moving across the camera. Furthermore, the advantage of MoG to other, less complex methods is that it recognises repeated movement (such as the branches of a tree moving in the wind) as background, not as a moving object. Competing techniques that might also be applicable here are one based on the codebook model [13], and one based on random aggregation [1].

3 Recommendations

3.1 Ideas and Concepts to Consider

3.1.1 Camera Placement

What the best portions of road are to monitor I do need to discuss, as Trafikverket has a lot more experience with this than I do. However, the direction from which the camera sees the road surface can have a large influence on the classification results. The camera placement more commonly seen in the literature cited above is such as to see a large portion of the road. This camera view yields a strong perspective in the image, as a portion of the road is close by, and a portion is far away. The major drawbacks of perspective is that different portions of the road are imaged at different resolution, meaning that possible texture on the road surface changes scale along the road. This change of scale makes it more difficult to find features that consistently characterise a certain road condition. Although it is possible to correct for the perspective view, it is not possible to correct for the reduced resolution: far away portions of the road would not show any texture at all. I would suggest hanging the camera over the road, looking straight down. This placement would minimise the perspective in the image. However, this placement is not compatible with the detection of specular reflection, if that is of interest.

3.1.2 Training and Testing Data

To build a system that autonomously determines the road condition, it is important to have lots of relevant image data. It is likely that a system trained with images from a single location will work best for that location, but also very unlikely to work well anywhere else (or even at the same location but looking in a different direction). The same can be said for time of day and the weather: if all training data is taken at noon on a sunny day, the system will not work very well at other times of day, at night, or on cloudy days. Therefore, it is important to collect data at geographically varied locations, at all times of day and night, and during different weather conditions and seasons. This is a difficult thing to do for a pilot study, of course. However, it should be possible to collect a small data set at a few locations and under various weather conditions, use that to train a pilot system, and while this system is in use, collect more data over time, improving the classifiers in the system as the data comes in. This would require a more complex module, but is presumably the best way of obtaining sufficient data for a reliable system. Furthermore, this method would allow each camera to train only on the images obtained by that same camera, and thus each camera would become more specialised to its own environment over time. Patrik Jonsson also suggested such a system during my interview with him.

3.1.3 Pre-Processing

A system to determine the road condition from image data would benefit immensely by the detection of passing cars and trucks. To do so, a method like MoG, as discussed in Section 2.2.3, should be applied. For this, it is necessary to have a continuous video feed; a single image every hour will not suffice. As transmitting a continuous video feed to a central location would consume too much bandwidth, I recommend applying this method in a locally installed computer. This computer would then send back to the central location either an image taken when no cars are detected, or a composite image taken over a short time span with all cars edited out. Optionally, this computer could run the complete classifier and simply send information on the road condition.

Patrik Jonsson mentioned that Trafikverket has developed a ruggedised computer for road-side installation. As none of the techniques discussed in this report are computationally expensive, such a computer would be able to do all the required computations. Note that the automatic toll system in Stockholm, developed by Q3, also performs all computations locally.

As I also discussed in Section 2.2.3, it would be important to mark, for each camera, which portions of the image represent the road, and which represent the shoulder, sky, trees, etc. Such a mask image could even exclude road markings (to avoid the issue found by Takeuchi et al. [24]). It is not clear whether information obtained outside the driving lanes will be relevant for prediction of conditions on the lanes, this should be investigated.

Further pre-processing should include determining the illumination condition (cloudy, sunny, dusk, night) and the direction of the illumination. This additional information would be added as features to the classifier, so that it can better interpret what is in the image. For example, a clear road in sunny weather and a snow-covered road at dusk might both be equally bright to the camera, unless there is an object of known reflectance in the field of view for comparison. Thus, taking the illumination condition into account will improve the comparison of intensities in the image. Likewise, the texture appearance of a rough surface such as a snow-covered road will change with the direction of illumination (as mentioned in Section 2.2.2), and thus a better comparison with the training set can be made if the direction of the illumination is taken into account. In lieu of the direction of illumination, the time of day could be used, as it would correspond with the position of the Sun.

3.1.4 Feature Extraction

I would suggest to explore the more modern texture features, such as the local binary patterns discussed in Section 2.2.2, besides the classical features such as intensity, gradients, and co-occurrence matrices. Which texture features will turn out to be relevant is difficult to predict, and will require a little experimentation. There are standard feature selection methods that can be used to discard features with little influence in the decision boundary. These features then do not need to be included in the final system. The AdaBoost classifier has feature selection implicitly built in.

3.1.5 Additional Features

Adding data other than image data is likely to make the classification more robust. As mentioned above, simple data such as the time of day would be very useful. And as shown by Jonsson [9, 10], including temperature, dew point, precipitation, and other weather information that is being collected currently, significantly improves the classification. Recording the sound from tires on the road, although very successful [18, 19], would only work if there is enough traffic driving by, and therefore probably not worth the effort.

As Patrik Jonsson extensively discussed in our interview (and reported in a paper [11]), it is possible to distinguish ice, snow and water by looking at a few infrared wavelengths. Infrared cameras are expensive, but Jonsson is developing a less expensive system that could be mounted next to each camera. The information of this infrared imager would supplement that of the video camera to further improve the reliability of the system.

3.1.6 Active Illumination and Light Polarisation

One of the more promising systems I found in the literature uses two cameras with orthogonally oriented polarising filters to see if the light reflected off the road is polarised or not. If polarised, the reflection is specular, indicating a smooth surface (either ice or water). This is a very good idea, but its dependence on the illumination makes it less reliable. If, however, there could be a strong light installed on the opposite side of the road (preferably near-infrared to avoid blinding drivers), the system would be completely independent of the natural illumination (and thus the weather and the time of day). The specular reflection can thus be detected even without polarising filters, and using a single camera. If there is no specular reflection, the intensity in the image would distinguish snow from clear road (given a fixed illumination, using the intensity in the image is quite trivial as well). I don't explore this option in more detail here because if falls outside the scope of the original request by Trafikverket.

3.1.7 System Ownership

As an academic, I have little knowledge of practises surrounding commercial system development. However, Patrik Jonsson had strong opinions about system ownership: If hiring a company to develop a road condition detection system, it is important that Trafikverket own the system, and keep the right to make changes to it at a later time. When buying a complete, closed system, it will not be possible to, e.g., update the cameras without buying a whole new system. If Trafikverket owns the system, it will also be easier, e.g., to add additional detectors to the system to try to improve its performance.

3.2 Proposed Plan for the Development of Image Analysis Software

Given that there are modern, very popular methods for pre-processing, feature extraction, feature selection and classification that have not yet been applied to the problem at hand (or at least not reported on in the scientific literature), I do think it makes sense to continue on to a phase II, in which these methods are tested and a computer program is developed that could be deployed.

However, I do strongly recommend to revise the current ideas at Trafikverket, as using only image data is likely to not produce very satisfying results. Adding additional weather information that is currently also collected together with the images will improve the results. As it is difficult to distinguish "black ice" visually, a camera-based system will also have trouble distinguishing this condition. Temperature alone will not always help distinguish ice from water, as it is not known how much salt is on the road, and even on unsalted roads there is an uncertainty around the 0°C mark [22]. Infrared imaging and measurement of reflection specularity seem to be good alternatives to solve these problems. I also strongly recommend the use of continuous video feed to be able to exclude moving objects (cars, falling snow close to the camera), which would require the use of a ruggedised computer next to each camera, and therefore the installation of additional hardware.

Nonetheless, there is no doubt that even a quite simple system will distinguish well between snow-covered and bare road surface, the minimal requirement set by Trafikverket for this project.

If Trafikverket decides to try alternative methods, the plan would be quite different. This assumes only image data and other currently collected additional data is used in the classification.

Requirements to start this phase II would be an appropriate data set (images with corresponding weather information, time, date, location, and known road condition), collected at several locations, at many different times of day and night, and in different seasons. Given such a data set, researchers appointed by Trafikverket can proceed to:

1. mark, for each location, the area in the image that corresponds to the driving lanes;
2. apply the various existing texture measures, and determine which ones provide the highest discriminating abilities between the various classes;
3. pick the most appropriate classifier according to the structure and number of the features to be used;
4. apply feature selection tools to see if any features are redundant;
5. thoroughly test the resulting classifier; and
6. devise a method to determine if something is obstructing the view of the road surface, so as to exclude that image or part of the image.

Items 1–5 might take four to eight weeks for an experienced researcher, or could be seen as a 20-week exjobb project with appropriate supervision. Time to complete item 6 is more problematic to estimate, as this is a much more difficult problem. If using continuous video feed, this problem does not need to be solved.

The next step after finishing items 1–6 above is to wrap all the algorithms into an independent program that can be deployed by Trafikverket. Many of the algorithms to be used exist as open-source code, and therefore might be used directly by Trafikverket (depending on license, of course). Some commercial packages could also be used to reduce the coding burden. As an academic, I have little experience in this, and therefore cannot estimate the time requirement. Furthermore, the deployed software will require maintenance, which, as an academic, I would not be able to provide. Therefore, I suggest contracting a company to provide the programming and maintenance, if Trafikverket cannot provide these services in house.

An additional system that could be build in, as discussed in Section 3.1.2, would allow the system to incrementally learn from the data recorded by each of the cameras, so that the system will become specialised for each location. In short, the system would be retrained with the newly obtained data periodically, and would require saving the data to be used for training. To build such a system requires somewhat complicated data management and user input, but the algorithms required are the same as that of the basic system developed in items 1–6.

References

[1] O. Barnich and M. Van Droogenbroeck. ViBE: A powerful random technique to estimate the background in video sequences. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 945–948, 2009. doi: 10.1109/ICASSP.2009.4959741.

[2] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[3] T. Bouwmans, F. El Baf, and B. Vachon. Background modeling using mixture of gaussians for foreground detection - a survey. *Recent Patents on Computer Science*, 1(3):219–237, 2008.

[4] Paul Bridge. Non-invasive road weather sensors. In *4th National Conference on Surface Transportation Weather (Indianapolis, Indiana)*, 2008.

[5] Feng Feng, Liping Fu, and Max S. Perchanok. Comparison of alternative models for road surface condition classification. In *the Transportation Research Board's 89th Annual Meeting*, pages 1–18, 2010.

[6] Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997. doi: 10.1006/jcss.1997.1504.

[7] Y. Horita, S. Kawai, T. Furukane, and K. Shibata. Efficient distinction of road surface conditions using surveillance camera images in night time. In *19th IEEE International Conference on Image Processing (ICIP)*, pages 485–488, 2012. doi: 10.1109/ICIP.2012.6466902.

[8] M. Jokela, M. Kutila, and Long Le. Road condition monitoring system based on a stereo camera. In *IEEE 5th International Conference*

on Intelligent Computer Communication and Processing (ICCP), pages 423–428, 2009. doi: 10.1109/ICCP.2009.5284724.

[9] P. Jonsson. Classification of road conditions: From camera images and weather data. In *IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA)*, pages 1–6, 2011. doi: 10.1109/CIMSA.2011.6059917.

[10] P. Jonsson. Road condition discrimination using weather data and camera images. In *14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 1616–1621, 2011. doi: 10.1109/ITSC.2011.6082921.

[11] P. Jonsson. Remote sensor for winter road surface status detection. In *IEEE Sensors*, pages 1285–1288, 2011. doi: 10.1109/ICSENS.2011.6127089.

[12] S. Kawai, K. Takeuchi, K. Shibata, and Y. Horita. A method to distinguish road surface conditions for car-mounted camera images at nighttime. In *12th International Conference on ITS Telecommunications (ITST)*, pages 668–672, 2012. doi: 10.1109/ITST.2012.6425265.

[13] Kyungnam Kim, Thanarat H. Chalidabhongse, David Harwood, and Larry Davis. Real-time foreground-background segmentation using codebook model. *Real-Time Imaging*, 11(3):172–185, 2005. doi: 10.1016/j.rti.2004.12.004.

[14] Wuttiwat Kongrattanaprasert, Hideyuki Nomura, and Tomoo Kamamura andKoji Ueda. Automatic detection of road surface states from tire noise using neural network analysis. In *Proceedings of 20th International Congress on Acoustics (ICA, Sydney, Australia)*, pages 1–4, 2010.

[15] Andreas Kuehnle and Wilco Burghout. Winter road condition recognition using video image classification. *Transportation Research Record: Journal of the Transportation Research Board*, 1627:29–33, 1998. doi: 10.3141/1627-05.

[16] M. Kutila, M. Jokela, J. Burgoa, A. Barsi, T. Lovas, and S. Zangherati. Optical roadstate monitoring for infrastructure-side co-operative traffic safety systems. In *IEEE Intelligent Vehicles Symposium*, pages 620–625, 2008. doi: 10.1109/IVS.2008.4621162.

[17] Matti Kutila, Maria Jokela, Bernd Roessler, and Jürgen Weingart. Utilization of optical road surface condition detection around intersections. In Gereon Meyer, Jürgen Valldorf, and Wolfgang Gessner, editors, *Advanced Microsystems for Automotive Applications*, VDI-Buch, pages 109–127. Springer, 2009. doi: 10.1007/978-3-642-00745-3_9.

[18] Kevin McFall. Artificial neural network technologies applied to road condition classification using acoustic signals. In *10th Standing International Road Weather Congress (Davos, Switzerland)*, pages 189–204, 2000.

[19] Kevin McFall and T. Niittula. Results of audio-visual winter road condition sensor prototype. In *11th Standing International Road Weather Congress (Sapporo, Japan)*, 2002.

[20] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 (7):971–987, 2002. doi: 10.1109/TPAMI.2002.1017623.

[21] R. Omer and Liping Fu. An automatic image recognition system for winter road surface condition classification. In *13th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 1375–1379, 2010. doi: 10.1109/ITSC.2010.5625290.

[22] Mats Riehm. *Measurements for winter road maintenance*. PhD thesis, KTH, Environmental Physics, 2012.

[23] Chris Stauffer and W. E. L. Grimson. Adaptive background mixture models for real-time tracking. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 246–252, 1999. doi: 10.1109/CVPR.1999.784637.

[24] K. Takeuchi, S. Kawai, K. Shibata, and Y. Horita. Distinction of winter road surface conditions using road surveillance camera. In *12th International Conference on ITS Telecommunications (ITST)*, pages 663–667, 2012. doi: 10.1109/ITST.2012.6425264.

[25] M. Yamada, K. Ueda, I. Horiba, and N. Sugie. Discrimination of the road condition toward understanding of vehicle driving environments. *IEEE Transactions on Intelligent Transportation Systems*, 2(1):26–31, 2001. doi: 10.1109/6979.911083.

[26] Muneo Yamada, Koji Ueda, Isao Horiba, Shin Yamamoto, and Sadayuki Tsugawa. Detection of wet-road conditions from images captured by a vehicle-mounted camera. *Journal of Robotics and Mechatronics*, 17:269–276, 2005.

[27] Chunyu Zhang, Wen Wang, Yong Su, and Xue Bai. Discrimination of highway snow condition with video monitor for safe driving environment. In *5th International Congress on Image and Signal Processing (CISP)*, pages 1241–1244, 2012. doi: 10.1109/CISP.2012.6469850.