Video-detection of winter road state

$200M annual cost of Wint. Maint.

100 cameras — RWIS sites

Andreas Kuehnle

Center for Research on Transportation and Society

Dalarna University

78188 Borlänge

Sweden

Abstract

This work examines the determination of winter road state from video images. The images may be either in black and white or color. The road state is one of five classes: dry, wet, snowy, icy, snowy with tracks. Image-related measures of the road state are calculated and given to a classifier, which determines the road state. The classifier, based on neural networks, is trained with examples, and based on these, is able to classify approximately 75% of unknown images correctly. There seems to be some advantage in using a color camera over a black and white camera.

Introduction and background

Winter weather can significantly affect daily life. Driving safety and mobility, in particular, are often compromised during adverse winter (road) conditions, with corollary economic and societal effects.

The goal of winter road maintenance is to alleviate the adverse effects of winter road conditions. The need for such maintenance measures is determined by observation of the road state, or projections of the road state in the near future. One form of observation is manual, on-site determination by an observer of the prevailing conditions. Manual observation of this sort is relatively expensive and not continuously available as observers sometimes need to travel to the area of interest first.

The goal of the work presented here is to automate the determination of road state by using intelligent, on-site video cameras. These cameras look at the road and give the image to a computer, which determines the road condition autonomously, having learned from examples how different sorts of road conditions appear. The system can generate information whenever requested or continuously, as desired. The goal with automatically generating this road state information is to aid decision making (through improved diagnosis) and to improve the quality control of maintenance measures (improving the cost effectiveness of the system).

There is only one directly related reference that we have found in this area (Chen, 1991). In this work, a laser illuminated the road surface, producing a black and white image which was analyzed. The goal of the study was to determine if black ice was present or not. The analysis used texture measures, and achieved a 94% rate of correct classification.

Indirectly related work includes the analysis of pavement distress from video images (Koutsopoulos et al, 1991) and the determination of skin cancer type from video images (Lee, 1994). All three of these references provided ideas for this work. The run length statistics of (Chen, 1991) and the color measures of (Lee, 1994) are used directly, while the crack (edge) statistics of (Koutsopolous et al,1991) provided inspiration for the edge-based features used here.

Method

The video-detection of winter road state relies on four steps. These steps are image acquistion, feature extraction, feature selection and classification with a neural network. These steps are described below.

1. Image acquistion is the first step. The existing network of roadside cameras is currently black and white, but will be upgraded to color cameras. We therefore looked at both black and white images and full color images in the hope that 1) either the existing network would suffice or that 2) more information and hence performance could be gained from color images.
2. Feature extraction calculates quantities - *features* - that describe the image in some way. This step generally reduces the amount of information needed for determination of the road

state (a picture has about a megabyte of information present, as opposed to a handful of features). This might involve calculating the average brightness of the image, for example, which is useful in determining whether there is snow on the road or not (snow on the road makes the image relatively bright). The features used here used only that portion of the image belonging to the road for the calculations and ignored what was around the road.

3. Feature selection chooses a subset of the total feature set which will be given to the classifier. This reduction serves to simplify the classifier, reduce the amount of examples required for training the classifier, and to (ideally) select those features containing the best information (it is difficult to know a priori which features are best). There can be many possible subsets if the number of features is large. Once the feature subset has been chosen, it is retained.
4. Classification builds (and then uses) a neural network that can classify unknown images. The neural network is trained with examples of images from each road state class, and learns how each road state appears (as expressed in its features). Once the training is complete, the system can be used to classify new images. The internal details of the neural network are varied during training to achieve the best results.

Details of each step are presented below. Fig. 1 shows a block diagram of the working system. Feature selection occurs during the training phase of the neural network.

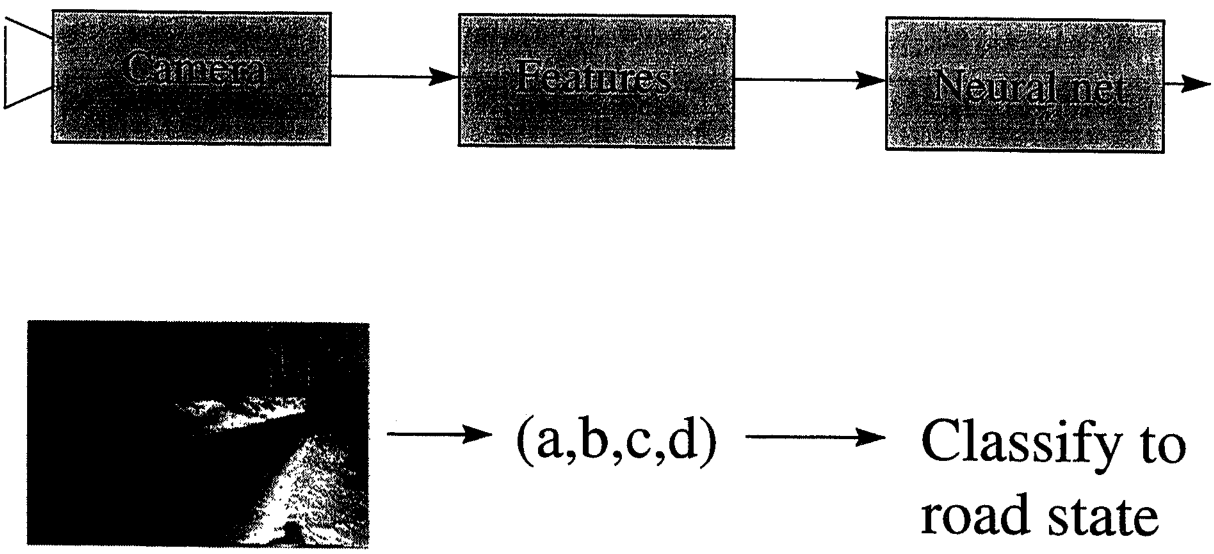


Fig. 1 Block diagram of image-based winter road state classification system.

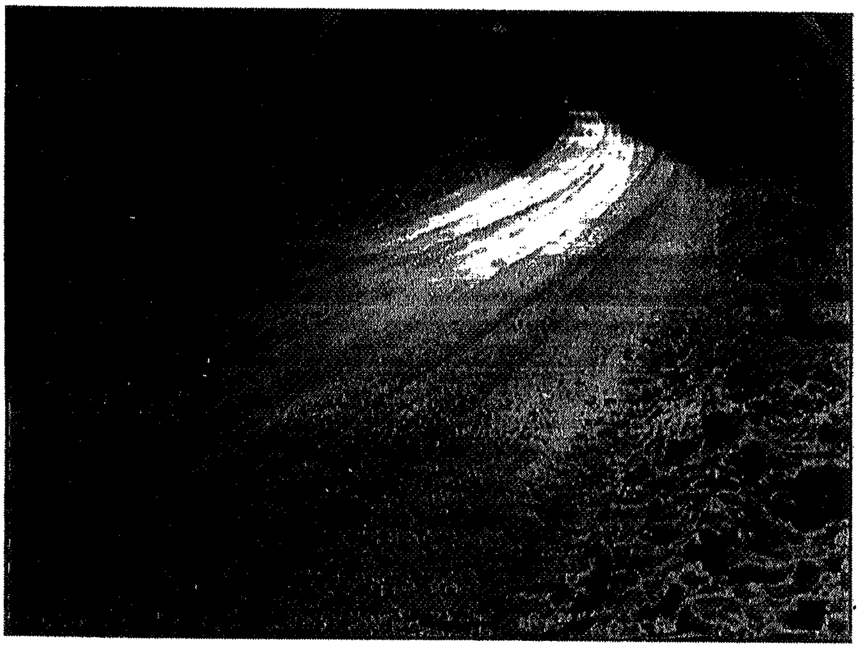
System details

This section presents the details of each of the above steps. As the system is not finalized, we simply describe the prototype that we implemented.

1. Image acquisition consisted of standing at the roadside and filming with a handheld color Sony Hi-8 camcorder. No consideration was made of the lighting, distance to the road, shadows, or other factors. The images were then digitized using the Y/C input on a Power Macintosh, and converting these to the TIFF file format (needed for further processing) with Adobe Photoshop. Black and white images were calculated by extracting the intensity value from the original RGB signal. Fig. 2 below shows a typical icy road scene.
2. Feature extraction extracted 34 features from the image. These include the average red, green, blue and gray (RGB/gray) levels, the variance of the RGB/gray levels, the coloredness for each color channel (calculated, for example, as red/(red+green+blue)), the average edge strength (calculated with the magnitude of the image when convolved with the Laplacian and Robert's masks (Schalkoff, 1989)), ratios of the color variances, run length distribution statistics (Chen, 1991), and the average closeness to white of the image (calculated with the average angle between a color pixel and white), and the predominant color in the image.
3. Feature selection extracted between three and five important features from the initial total set of 34. The selection was done in two stages. A subset of the total feature set (34) was first selected on the basis of the Mahanalobis distance (Bishop, 1995) that these produced between classes. Those 8 to 14 features that produced the largest average interclass separation, the largest minimum interclass separation and the largest maximum interclass separation were retained. Second, a brute force search was then used to iteratively present every single (three to five element) feature combination to a neural network, whose performance determined which feature combinations were finally best. As neural networks (generally) start from a random point during training, each feature combination was used 25 times for training . A total of 400 epochs per training was used. The performance of the neural network was determined by testing it with images it had not been trained with.
4. Feature selection already incorporated some network training. Those feature sets which produced good performance were further investigated by more training, as well as variations of the network architecture. The network architecture was of two types (logsig input layer, logsig hidden layer, linear output layer and logsig transfer functions in each

layer). The number of input neurons was the number of features. The number of hidden neurons varied between 3 and 8. The number of output neurons was either one (witn target class values from 1 to 5 and a linear output neuron, or target class values between 0 and 1 and a logsig output neuron) or 5 (and a target vector, for example for class 2, of {0 1 0 0 0}). 48 training samples were used, as well as 21 test samples. Early stopping was used.

Fig. 2 Example of icy road scene used in study.



Results

The results consist of two types: how well the classifier performs with unknown images (that is, what percentage of unknown images classifies correctly) and how classes classify (that is, are mistakes made with all classes, or do some classes classify correctly while others do not). Table 1 presents various feature set types, the members of these feature sets, and the classification performance of the neural network. The neural network had three hidden neurons (increasing the number of hidden neurons worsened the performance) and one linear output neuron (multiple output neurons did not help, nor did a logsig output neuron).

Table 1 Feature sets and the resulting classifier performance.

Feature set type	Feature set members	Network performance
3 feature black and white	average edge strength (Roberts filter), short run emphasis (vertically), and run length distribution (horizontally)	71%
4 feature black and white	mean gray, gray variance, short run emphasis (vertically) and run length distribution (horizontally)	71%
3 feature full color	green variance, ratio of red variance to blue variance, run length distribution (vertically)	62%
4 feature full color	gray variance, ratio of red variance to blue variance, mean green, run length distribution (horizontally)	76%
5 feature full color	mean value of green, mean gray value, green variance, gray variance, average edge strength (Roberts)	62%

As can be seen, rates of correct classification are typically between approximately 60 and 80%. Color improves performance a little (though the relatively small amount of data precludes making a definitive statement), though the addition of more features does not help.

Table 2 presents how the different classes classify. The top row shows the number of members of each class in the test set (thus "5 dry roads" means that the test set had five samples of dry roads). The leftmost column shows the feature set type. The inner cells on the table shows now that column (class) classifies. Errors are shown in italics. The members of the feature sets are listed in Table 1.

Table 2 Classification of test images, with misclassifications in *italics*.

Feature set	5 dry roads	5 wet roads	3 track snow	3 snowy roads	5 icy roads
3 feature B/W	5 dry	4 wet, *1 dry*	*3 snow*	3 snow	3 icy, *2 wet*
4 feature B/W	5 dry	4 wet, *1 snow*	*3 snow*	3 snow	3 icy, *1 dry, 1 snow*
3 feature color	4 dry, *1 undefined output*	3 wet, *2 icy*	*2 icy, 1 snow*	3 snow	3 icy, *1 dry, 1 wet*
4 feature color	5 dry	5 wet	2 track snow, *1 snow*	3 snow	1 icy, *4 wet*
5 feature color	4 dry, *1 icy*	3 wet, *2 icy*	*3 snow*	3 snow	3 icy, *2 wet*

As can be seen, dry, wet and snowy roads generally classify correctly. Roads with tracked snow often misclassify as snow or ice, but more seriously, icy roads often misclassify to wet or other classes. The neural network thus does not distribute its mistakes evenly, but concentrates them on the icy and tracked snow classes.

Discussion

The prototype system described here classifies approximately ¾ of all unknown images correctly. Chance should give a correct classification rate of one/number of classes, or 20%. While the performance here is thus better than pure chance, we still do not find it acceptable, or the best that we can achieve.

Three major improvements can be made to the system. These are:

1. collect more data
2. develop new features
3. refine the neural network classifier

Collecting more data will give us a more firm basis on which to describe how various road states appear. We are currently increasing the database to a few hundred images. A small number of samples (69 were used here) simply cannot adequately represent a whole population. We will also examine the effects of illumination, particularly at night, where infrared lamps replace the sun.

Developing new features will let us tune the system better, in particular to allow distinguishing between tracked snow and other classes, and ice and other classes. An edge filter sensitive to tracks approximately parallel to the road may be useful for tracked snow. Tuning the run length calculations (there are arbitrary parameters involved) may allow better recognition of ice.

Refining the neural network and a deeper analysis of the feature clustering should improve the overall system perfomance. In particular, weighting misclassifications of icy scenes may be useful to eliminate this problem.

We hope to achieve a better than 80% rate of correct classification given these improvements, and also hope that mobile cameras might be useful as input to the system (thus providing a larger degree of network coverage).

Acknowledgements

The suggestions of Corinne Ledoux regarding network architecture and training with early stopping are appreciated, as is the pointer of Sten Sundin to the work on skin cancer diagnosis.

References

Bishop, C.(1995) *Neural Networks for Pattern Recognition*, Oxford, 1995.

Chen, Y. (1991) ,The use of image analysis for black ice detection, University of California, Berkeley, in *Recherche, Transports, Securite*, English Issue, No. 7.

Koutsopoulos, H. and El Sanhouri (1991), I. Methods and algorithms for automated analysis of pavement images. In *Transportation Research Record 1131*, pp. 103-111.

Lee, H.C. (1994) Skin cancer diagnosis using hierarchical neural networks and fuzzy logic, Ph.D. Thesis in computer science at University of Missouri-Rolla, located on World Wide Web at http://www.cs.umr.edu/hlee/thesis/htesis.html.

Schalkoff, R. (1989) *Digital image processing and computer vision*. Wiley and sons, New York.