2013 IEEE/RSJ International Conference on
Intelligent Robots and Systems (IROS)
November 3-7, 2013. Tokyo, Japan

A Roadwork Scene Signature Based On the Opponent Colour Model

Bonolo Mathibela, Ingmar Posner, and Paul Newman

Abstract— **The presence of roadworks greatly affects the validity of prior maps used for navigation by autonomous vehicles. This paper addresses the problem of quickly and robustly assessing the gist of traffic scenes for whether roadworks might be present. Without explicitly modelling individual roadwork indicators such as traffic cones, construction barriers or traffic signs, our method instead only exploits the engineered visual saliency of such objects. We draw inspiration from opponent colour vision in humans to formulate a novel roadwork scene signature based on an opponent spatial prior combined with gradient information. Finally, we apply our roadwork scene signature to the task of roadwork scene recognition, within a classification framework based on soft assignment vectorization and RUSBoost. We evaluate our roadwork signature on real life data from our autonomous vehicle.**

I. Introduction

Roadworks visually demarcate the presence and extent of road maintenance, construction and accident scenes. The layout of a roadwork scene depends on the underlying cause of the disruption, which in turn influences the objects used to denote the scene. For example, highway maintenance is typically denoted by traffic cones and traffic signs, while city works are denoted by traffic signs, construction barriers and traffic cones. Directly modelling a roadwork scene requires a formulation for the relationships between these different objects within the scene. A naive approach is to detect individual objects in the scene and classify a scene based on occurrences of roadwork objects. Such an approach is both computationally expensive and requires training multiple object detectors. Our more sophisticated model however, describes the entire roadwork scene using a global signature without explicitly modelling individual objects within the scene. This approach is inspired by the relationship between opponent colour vision and the design of roadwork objects.

Although roadworks are temporary, their frequency of occurrence is surprisingly high. In 2009/2010 alone, Transport for London reported an estimated 370 000 roadworks [1], a figure that is typical for most major cities. Roadworks introduce new infrastructure on the road, thus dramatically changing its layout and negatively affecting the accuracy of prior maps [2] which have recently been exploited to constrain tasks such as traffic light detection and mapping [3], [4]. Such prior maps can no longer be trusted at roadwork sites and autonomy should potentially be revoked in a safe manner. There therefore exists a need for an efficient mechanism to quickly recognise roadwork scenes in order to update prior assumptions about the world and offer manual control to the human driver (or alternative navigation algorithm) as shown by Figure 1- this motivates our work.

Fig. 1. Our proposed autonomous vehicle interface: when roadworks are detected (top), manual control is offered to the driver (bottom). The roadworks shown here were successfully recognised using our roadwork scene signature. Such a system has two important applications: offering manual vehicle control when autonomy is limited and cuing the automatic switch from a normal to a roadwork autonomous navigation algorithm.

The main contribution of this work is an elegant and robust framework for roadwork scene recognition based on an opponent colour roadwork scene signature. Following a review of related works in the next section, we discuss the opponent colour model in Section III. Our roadwork scene signature is introduced in Section IV followed by an evaluation in Section V. Finally, we conclude in Section VI.

II. Related Works

Colour information plays a vital role in object and scene recognition, and many colour descriptors have been proposed over the years (see [5] for a good overview). The Hue Saturation Value (HSV) colour space has gained popularity within the vision community where it has been applied to a large corpus of problems such as face detection [6], traffic sign recognition [7], and pedestrian detection [8].

Scene recognition is often posed as a similar image retrieval problem by the computer vision community. This approach requires a representation of the image as a fixed-size vector created from quantised descriptors. Bosch et. al. apply probabilistic Latent Semantic Analysis (pLSA) to a bag of words representation of each image to create a descriptor. A multiway classifier is then trained on the topic distribution of each image [9]. In similar work, Siagian and Itti [10] use a multi-scale set of early-visual features to represent an image as a low-dimensional signature vector. In scene classification literature, contextual information is often captured by modelling spatial dependencies between classified segmented regions of the image to aid scene recognition [11]. Our method differs from previous work in that spatial dependencies between regions in the image are captured by an opponent colour spatial prior (see Figures 5 and 8). Furthermore, we extend the Bag-of-words approach to implement soft assignment vectorization based on standard deviations from codebook centroids. This allows us to capture the degree of similarity between features.

Other colour spaces have been used before: Bratkova et al. present a practical opponent colour space, known as oRGB, useful for graphics applications such as colour adjustment, transformation, and transfer [12]. Ishizuka and Hirai segment red road signs using the HSV colour space and later perform traffic sign recognition based on opponent colour filters [13]. Other colour descriptors such as colour GIST [14] and colour SIFT [5] have also been proposed before. The most relevant related work is the Bag-of-colors (BOC) descriptor proposed by Wengert et. al. [15] for improved image search. This descriptor is based on the generation of a CIE Lab colour dictionary (codebook) which is used to represent an image as a normalised histogram colour signature.

Previous work on roadworks is limited to the automatic generation of motorway roadwork maps for driver assistance systems [16], and the detection and classification of motorway concrete and steel roadwork safely barriers to generate maps [17]. In our previous work [2], we considered anticipation of roadworks by combining online city council location priors with local observations to allow us to accurately map roadwork layouts and discount inaccurate prior information.

In formulating our roadwork scene signature, we do not detect individual objects in the scene but instead represent the scene using our roadwork signature. To the best of our knowledge, an opponent colour based roadwork scene signature has not previously been proposed for roadwork scene recognition in autonomous driving applications.

III. Roadwork Scenes and Opponent Colour

A. Roadwork Scenes

Roadwork objects are predominantly coloured red, green, yellow or blue (or some closely related hue) as seen in Figure 2. Concretely, roadworks are purposely designed to be salient to the human visual processing system, which perceives colour in an opponent-like fashion. This observation underpins our formulation of a roadwork scene signature.

Fig. 2. Typical roadwork scenes found in our datasets. The dominant co-occurrances of red-green and blue-yellow hues in the roadwork scenes, inspire our opponent colour formulation of a roadwork scene signature.

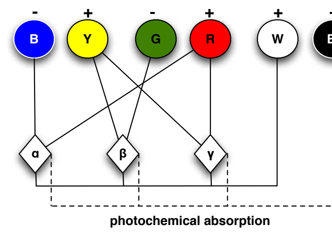

Fig. 3. A simplified schematic of the Hering opponent-colour model from [20]. The polarities indicate two possible modes of responses that are physiologically opponent in nature. The three pairs B-Y, G-R, and W-Bl are antagonistic (i.e. the concept of redish-green is foreign to the human visual system, whereas yellowish-red is perceived as orange). Here, α, β, γ are photosensitive biological retina materials aiding the formation of colour.

B. The Opponent Colour Model

The perception of colour depends on three factors: the illumination source, the material of the object being illuminated and the observer (human eye or camera) [18]. The back of the human eye has a photoreceptive region, known as the retina, containing rod and cone photoreceptors that allow humans to perceive monochromatic and colour vision respectively. Several visual processes have been proposed to explain the perception of colour in humans. Colour vision, while not studied in robotics, has widely been studied by the neuroscience and psychology community. In general, there exists two main theoretical accounts for the formation of colour on the human visual cortex: the Young-Helmholtz and the Hering theories [19]. The former model is a theory of trichromatic colour vision based on three primary colours (red, green, blue) and their combinations on the eye's cone photoreceptors to produce the full spectrum of colours perceived by humans. In contrast, Hering's theory postulates opponent colour vision involving three types of paired visual processes occurring in the visual cortex, such that each process has two opponent members. These sensory processes are antagonistic within each channel [20] and are summarised graphically by Figure 3.

In this paper, we limit our discussion of the opponent colour model to its effect on the design and manufacture of roadwork objects and how this knowledge can be exploited when formulating a roadwork scene signature. Figure 3 shows neuron colour responses of the visual cortex which are grouped into three antagonistic processes: yellow-blue (Y-B), red-green (R-G), and white-black(W-Bl) intensity. These responses are linked to three photosensitive biological

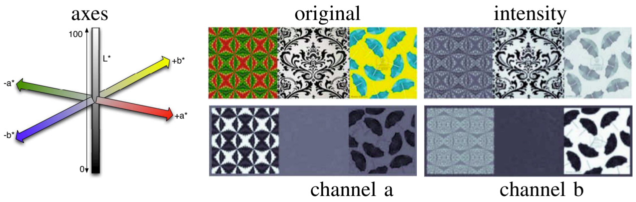

Fig. 4. Axes of the CIE Lab colour space are blue-yellow and red-green (both chromatic), and black-white (achromatic) occurring antagonistically. The red-green pattern is best highlighted by channel a, with the opponent colour intensities at opposite extremes. Similarly, the blue-yellow pattern is best highlighted by channel b. The intensity channel is also shown.

materials found in the eyes (denoted by α, β, γ). Hurvich and Jameson [20] showed that each opponent colour channel can be derived according to:

$$\begin{aligned} C_{y-b} &= k_1(\beta + \gamma - 2\alpha) \\ C_{r-g} &= k_2(\alpha + \gamma - 2\beta) \\ C_{w-bl} &= k_3(\alpha + \gamma + \beta) - k_4(\alpha + \beta + \gamma) \end{aligned} \tag{1}$$

where C_{x-y} represents an opponent colour channel and each k_i is a physiological constant.

In practice, the opponent colour model implies that certain combinations of colours cannot be perceived together. For example, blueish-green or yellowish-red describe well known colours cyan and orange. The concept of blueish-yellow or redish-green is however foreign to the human perception system. In Figure 2, we observe red-green roadwork barriers, blue traffic signs, and bright orange traffic cones and barriers. Roadworks are therefore designed to be easily perceived by the opponent colour process of the human visual cortex.

For our application, we seek an equivalent computer vision colour model that interprets red-green and blue-yellow antagonistically, and therefore perceives roadworks the same way a human does. In computer vision, the CIE Lab colour space is an opponent space that models perceptually uniform human vision by approximating the logarithmic response functions of the eye. As in human vision, the CIE Lab colour space has an achromatic channel (intensity) and two chromatic red-green and blue-yellow colour channels. Perceptual uniformity implies that changes to any of the Lab components induces similar changes in the *perceived colour*. Converting from classic RGB to CIE Lab involves a nonlinear transform. This opponent colour model is therefore useful for pattern recognition because it naturally segments patterns into perceptually meaningful opposing colour channels. This effect is illustrated by the righthand Figure 4 where the red-green and yellow-blue patterns are highlighted by the opponent colour transformation. This automatic pixel labelling characteristic is computationally efficient and a natural property of the colour model. Importantly, the CIE Lab colour space is consistent with the Hering colour model used to explain human vision and each channel encodes exactly two colours in an opposing fashion. Next, we formulate a roadwork scene signature inspired by these two properties.

IV. Roadwork Scene Signature

A roadwork scene signature is a compact image representation useful for characterising a road scene. We formulate this as a global signature based on the opponent colour model described in the previous section.

A. Opponent Based spatial Prior

In Section III, we showed the relationship between vision in humans and the design of roadwork objects. Furthermore, we discussed why the CIE Lab opponent colour space is a good approximation of human opponent vision for the task of roadwork scene recognition. A good roadwork scene signature should be discriminative and robust to scene clutter, which motivates the first step of the signature creation process: transforming the image to CIE Lab colour space. This transformation allows us to exploit the automatic pixel labelling properties of the opponent colour space in order to implicitly derive a spatial prior for the location of roadwork objects in a given image. Only the chromatic CIE Lab image channels (a and b) are considered in this formulation. Figure 4 shows that, for each of the chromatic channels, the opponent colours are at opposite intensity extremes (yellow is positive and blue is negative, similarly red is positive and green is negative). The peak gradient information is therefore concentrated at the opponent extreme regions where an image keypoint detector peaks. We exploit this for efficiently extracting descriptors only at relevant image locations. This computationally efficient spatial prior allows us to avoid naively detecting individual objects in a given scene (see Figure 5). Our formulation differs from traditional approaches where many keypoints are used and often densely sampled.

B. Image Descriptor Extraction

Following the opponent based spatial prior described in the previous section, we extract a descriptor at each detected image keypoint. For a given colour channel and detected keypoint, a 128 bit SURF descriptor [21] is extracted as shown in Figure 6. We choose the SURF descriptor for its speed and robustness [21]. Other descriptors such as RGB histograms, SIFT [22], colour moments or moment invariants are also valid possible candidates for our proposed framework, however a comparison of different descriptors is beyond the intended message of this paper, and we refer the reader to [5] for a good comparison.

Our proposed descriptor differs from the Bag-of-Colors model [15] even though both methods exploit the CIE Lab colour space. Firstly, SURF keypoints are only detected at points most likely to contain roadworks due to our initial spatial prior. In contrast, the Bag-of-colors approach creates a colour descriptor by considering every pixel in the image to increment a k_c-dimensional vector, where k_c is the colour codebook length. Secondly, in addition to colour, our descriptor also captures texture information.

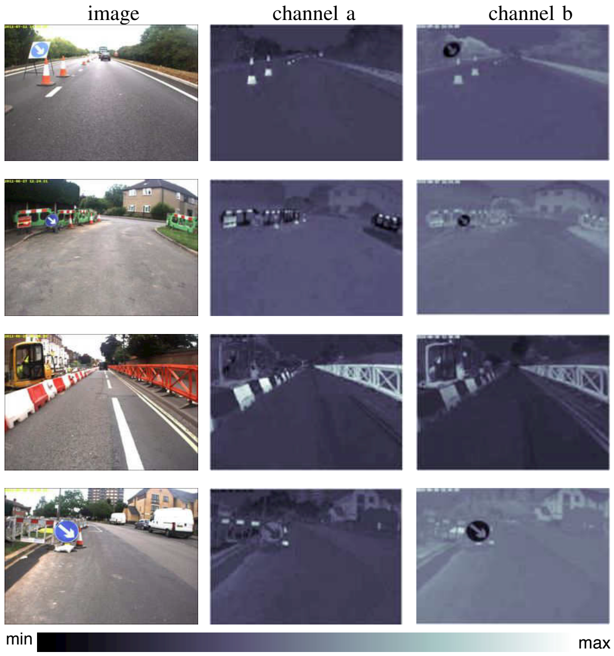

Fig. 5. Results from our opponent based spatial prior showing the automatic labelling of pixels in the a and b CIE Lab colour channels. The intensity scale is shown at the bottom of the figure. Notice how the blue traffic signs are well segmented in the b (blue-yellow) channel. Similarly, the barriers and traffic cones are well segmented by the a (red-green) channel.

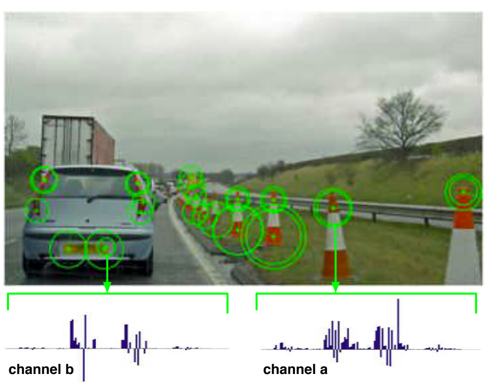

Fig. 6. Our image descriptor comprises of a 128 bit SURF descriptor [21] computed at each keypoint in the image for both opponent chromatic colour channels. The circles denote the scale and orientation of a descriptor.

C. Codebook Generation and Vectorization

For a given image, the number of extracted descriptors depends on the number of detected keypoints. This number varies with each image, thus a method to represent every image as a single vector is required. To achieve this, a codebook is generated and each image is vector quantised against this codebook in order to derive a single vector representation. Our codebook generation and vectorization procedure is an improved version of the Bag-of-Words approach for scene recognition [23] and is summarised by Figure 7. Codebook generation proceeds as follows:

1) Collect training images containing road scenes (roadwork and non-roadwork)
2) Create a codebook: for all training images extract image descriptors according to Section IV-B. K-means cluster these to create a codebook vocabulary, choosing k based on precision recall performance on a holdout dataset. Each cluster centroid, c_i then represents a visual word in this codebook.

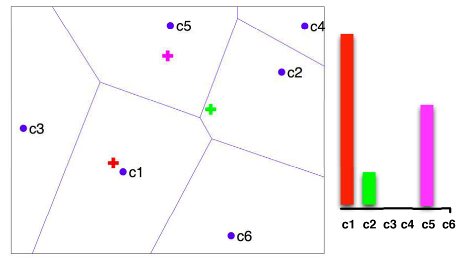

Fig. 7. Soft assignment vectorization for a codebook $C = \{c_i\}_{i=1}^{6}$ with six visual word centroids . Three descriptors (shown as crosses) were detected in the image. The histogram on the right shows the final vectorized image representation. Descriptors closest to the centroid receive higher weighting in the vectorized histogram representation (Equation 3). In this way, we capture the distribution of the descriptors with respect to the nearest neighbour centroid. This exploits the perceptual uniformity (colour differencing) property of CIE Lab.

Given a codebook, we can now vectorize a new image by quantising it against the codebook. The Bag-of-Words model represents each image as a histogram vector of length C. Using a hard assignment rule, each histogram bin is updated by one when a detected descriptor is closest to its associated visual word. Thus, given descriptor x and nearest neighbour $NN(x) = c_i$ the i^{th} histogram bin h_i is updated according to:

$$h_i = h_i + 1 \ \ if \ NN(x) = c_i. \quad (2)$$

In our work, we adopt a soft assignment approach instead. We update the histogram bin by the inverse standard deviation from the closest visual word according to:

$$h_i = h_i + \left[\sqrt{\frac{1}{\mid m \mid} \sum_{m} (x_m - c_{im})^2} + \varepsilon \right]^{-1} \ \ if \ NN(x) = c_i \quad (3)$$

where $m = \{1, 2, \ldots, k\}$ indexes the elements of a visual word (or descriptor) and ε is a small number ensuring that the denominator is nonzero if the descriptor and codebook visual word match exactly. This soft assignment approach captures the degree of similarity between a descriptor and its closest visual word and indirectly models the feature distribution. We can adopt this approach because the perceptually uniform CIE Lab colour model represents a non linear transformation into a new space that is better able to predict visual colour difference between two points using the Euclidean distance metric [24]. Furthermore, van Gemert et. al showed that explicitly modelling visual word assignment decreases classification ambiguity when compared to the traditional hard assignment model [25]. Many soft assignment schemes exist, however our approach is inspired by work done by Jegou et. al. [26] to aggregate local descriptors into a single image representation for image search: a difference between

no spacial prior | with opponent prior

Fig. 8. Visual words for roadwork scenes found in our datasets. The SURF threshold remains constant throughout. Significantly fewer visual words are extracted by incorporating the opponent colour model (right). These are also extracted in regions most likely to contain roadwork objects.

a descriptor and visual word is calculated then dimensionality reduction is done to represent the image using a single vector.

The key advantage of our formulation is that the opponent colour based spatial prior directs our descriptor extraction algorithm to regions in the image most likely to contain roadworks because hues not encoded by the opponent colour channels are suppressed. This phenomenon is demonstrated by Figure 8 which shows visual words detected with and without our spatial prior. Our roadwork signature, h therefore comprises of fewer but relevant visual words resulting in a compact vectorized image representation.

D. Training and Validation

The codebook generation and vectorization scheme described in Section IV-C results in a roadwork scene signature (feature vector) for a given image. Given a set of training images (containing non-roadwork and roadwork scenes) we extract roadwork scene signatures for all images and use these input features to train a classifier for roadwork scene recognition. Our classifier choice is influenced by the fact that, although roadworks are common [1], they are finite and occupy limited sections of the road. This results in a skewed number of examples for roadwork versus non-roadwork classes. The RUSBoost algorithm [27] handles class imbalance by combining random under sampling (RUS) with AdaBoost [28] resulting in an improved model of the minority class by removing majority class examples. We therefore train a RUSBoost classifier using tree learners and 500 rounds within a five-fold cross validation scheme.

	Source	Description	# Pos e.g.	# Neg e.g.
Train	Geograph	Codebook	536	1002
	Geograph	Classifier	1032	2042
Test	Town	Dataset 1	57	89
	Residential	Dataset 2	60	139
	Motorway	Dataset 3	71	28
	City	Dataset 4	1443	2075
	Motorway + city	Dataset 5	1195	704

TABLE I

A SUMMARY OF THE DATASETS USED FOR TRAINING AND EVALUATION

V. Experimental Results

Our roadwork scene signature was evaluated on data collected over several months using an autonomous vehicle. The training data was sourced online from the Geograph Project database[1] for Britain and Ireland [29] containing images submitted by the public. Table I describes our datasets.

To evaluate the effectiveness of our spatial prior, we compare the number of SURF keypoints extracted in 100 images with and without the opponent spatial prior. From Figure 9, it is clear that the number of SURF keypoints is significantly reduced when using our opponent colour spatial prior formulation without compromising on accuracy, since features are extracted only at relevant locations (see Figure 8 for examples). Furthermore, Figure 10 shows that, for our opponent colour based formulation, the number of SURF keypoints is in general smaller for images not containing roadworks which confirms our intuition about the visual saliency of roadworks in images. In other words, roadworks are designed to be detected by the human visual system which operates in an opponent like fashion (see Section III for a discussion). To understand this further, we computed the Bayes optimal threshold for minimising the misclassification error based on the number of SURF keypoints detected in a given image. A misclassification rate of 30% was achieved with a simple feature count threshold classifier.

An important design consideration for our roadwork scene signature is how to select an appropriate codebook size, k. Care should be taken not to overfit the data whilst ensuring that a representative and discriminative number of visual

[1] Creative Commons License

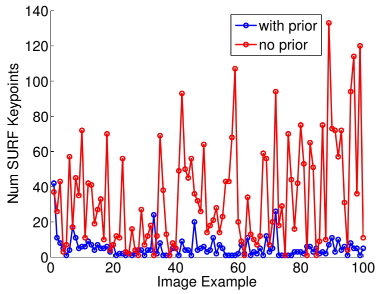

Fig. 9. The number of SURF keypoints detected without using the opponent colour spatial prior is significantly higher than when the transform is applied. Although our roadwork scene descriptor uses fewer keypoints, the descriptors extracted at these keypoints are the most relevant ones, as illustrated by Figure 8. Accuracy is therefore not compromised.

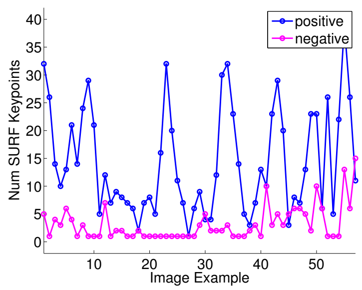

Fig. 10. In general, fewer keypoints are detected for negative examples which demonstrates the visual saliency of roadworks i.e. our descriptor focuses on roadwork objects highlighted by our spatial prior, while suppressing non-opponent noise.

words is used. We experimentally determined the value of k by computing the precision and recall on a holdout dataset for various values of k, and found $k = 100$ to give the best performance. Our codebook size is indeed very small, which is again a property of the spatial prior discussed in Section IV-A and demonstrated by Figure 5.

In order to test the effectiveness of our roadwork scene signature, we apply it to the complex task of roadwork scene recognition. We also deliberately trained and tested on completely different datasets (online sourced training set and test set from autonomous vehicle) in order to test the generalisation and robustness of our system. For the test datasets collected by our autonomous vehicle, we compute Receiver Operation Characteristics (ROC curves) to illustrate the performance of our classifier with varying RUSBooost discrimination threshold (see Figure 12 left). The ROC curves show good performance illustrating the discriminative property of our roadwork scene signature. Figure 1 shows examples of successful roadwork detections across all datasets using our roadwork scene signature and the proposed associated user interface in our autonomous vehicle. We also tested our system using a naive SURF descriptor on grey level images keeping all other experiment parameters the same. Figure 12 (right) shows the results demonstrating significantly reduced performance in terms of the ROC curves. This illustrates the significance of our opponent colour vision spatial prior.

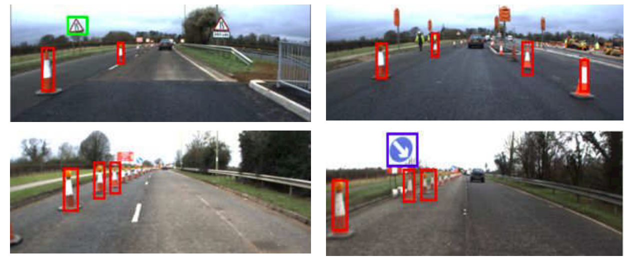

Fig. 11. Some examples of traffic cone and keep right sign detector outputs.

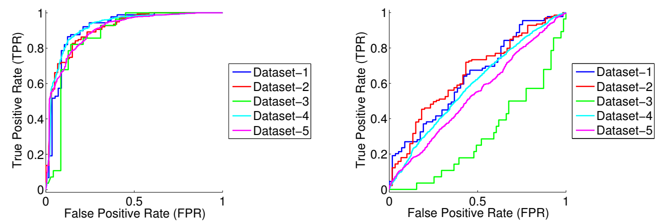

Fig. 12. LEFT: ROC curves showing the TPR and FPR achieved for the datasets described in Table I. Each curve represents a different dataset collected using our autonomous vehicle. Good performance is achieved across all datasets using our roadwork scene signature. RIGHT: Recognition performance is significantly reduced without using our opponent colour spatial prior (keeping all other experiment parameters the same).

To the best of our knowledge, there is no similar roadwork scene recognition work to compare against, however we implement a naive alternative to our method based on object recognition for comparison. To do this, we train traffic cone and keep right sign classifiers (see Figure 11) using dictionary template matching features [30] trained within a GentleBoost framework [31]. An image is then classified based on the presence of roadwork objects in the scene.

For this comparison, we assume that the occurrence of traffic cones in a given scene is conditionally independent of the presence of keep right traffic signs given the class label. Each image is therefore represented by Naive Bayes: $p(C|f_1, f_2) = \frac{1}{Z}p(C)p(f_1, f_2|C)$, where the features, f_1 and f_2 denote the presence of traffic cones and keep right signs respectively, Z is the normalisation constant, and classes C are roadwork and non-roadwork. We use a maximum likelihood estimate of the parameters. Results of our comparison are shown in Figure 13 demonstrating that our roadwork scene signature outperforms the object detection approach. The reasons for this are twofold: firstly, the presence of multiple traffic cones (or keep right signs) does not necessarily imply a roadwork scene since keep right signs are not unique to roadwork scenes and traffic cones can appear for a number of reasons such as marking off where children are playing or advance warning of hazards. Secondly, a roadwork scene recognition system based on individual detectors is sensitive to detector performance. Contrastingly, our proposed roadwork scene signature does not require the detection of individual roadwork objects in the scene and is hence faster to train, apply to an image and has the added advantage of being more accurate.

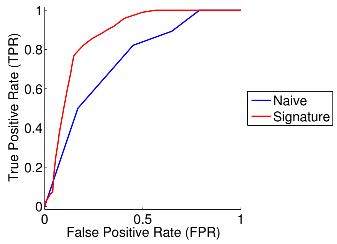

Fig. 13. A comparison showing ROC performance using our roadwork scene signature compared to roadwork object detection within a Naive Bayes scene recognition framework. Our method outperforms the naive approach and has the added advantage of not explicitly modelling individual roadwork objects in the scene.

VI. Conclusions and Future Work

In this paper, we demonstrate that roadworks are purposely engineered to be salient to the human opponent colour based visual system. We use this fact to formulate a spatial prior for the location of roadwork objects in a given image, based on the CIE Lab opponent colour space (to approximate human vision). We then derive a unique roadwork scene signature, a global image representation that captures the gist of roadwork scenes without explicitly modelling individual objects. We demonstrate the usefulness of this roadwork signature for the task of roadwork scene recognition for an autonomous vehicle. Our system is evaluated on real data from an autonomous vehicle and is shown to perform well across multiple datasets containing varied types of roadwork images. Future work includes adding a temporal dimension to our framework and testing our system on video sequences. We will use this roadwork scene signature as a prior indicator when mapping out drivable regions at roadwork sites.

VII. Acknowledgment

Thanks to Barry Houter for the Geograph dataset and the Rhodes Trust and National Research Foundation for funding. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

References

[1] G. Emmerson, "Minimising disruption from roadworks," 2010.

[2] B. Mathibela, M. A. Osborne, I. Posner, and P. Newman, "Can priors be trusted? learning to anticipate roadworks," in *15th Intl. IEEE Conference on Intelligent Transportation Systems (ITSC 2012)*, Anchorage, AK, USA, 2012.

[3] N. Fairfield and C. Urmson, "Traffic light mapping and detection," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, may 2011, pp. 5421 –5426.

[4] J. Levinson, J. Askeland, J. Dolson, and S. Thrun, "Traffic light mapping, localization, and state detection for autonomous vehicles," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, may 2011, pp. 5784 –5791.

[5] K. E. van de Sande, T. Gevers, and C. G. Snoek, "Evaluating color descriptors for object and scene recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1582–1596, 2010.

[6] O. Ikeda, "Segmentation of faces in video footage using hsv color for face detection and image retrieval," in *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, vol. 3, sept. 2003, pp. III – 913–6 vol.2.

[7] W. Shadeed, "Road traffic sign detection in color images," in *IEEE International Conference on Electronics, Circuits and Systems*, 2003.

[8] S. Walk, N. Majer, K. Schindler, and B. Schiele, "New features and insights for pedestrian detection," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE. San Francisco: IEEE, 2010.

[9] A. Z. Anna Bosch and X. Munoz, "Scene classification using a hybrid generative/discriminative approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, pp. 712–727, April 2008.

[10] C. Siagian and L. Itti, "Rapid biologically-inspired scene classification using features shared with visual attention," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 2, Feb 2007.

[11] S. Kumar and M. Hebert, "Discriminative random fields," *International Journal of Computer Vision*, vol. 68, no. 2, pp. 179–201, 2006.

[12] M. Bratkova, S. Boulos, and P. Shirley, "orgb: A practical opponent color space for computer graphics," *Computer Graphics and Applications, IEEE*, vol. 29, no. 1, pp. 42 –55, jan.-feb. 2009.

[13] Y. Ishizuka, Y. Ishizuka, and Y. Hirai, "Segmentation of road sign symbols using opponent-color filters."

[14] A. Oliva and A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," *International Journal of Computer Vision*, vol. 42, no. 3, pp. 145–175, 2001.

[15] C. Wengert, M. Douze, and H. Jégou, "Bag-of-colors for improved image search," in *ACM Multimedia*, Scottsdale, United States, Oct. 2011, qUAERO. [Online]. Available: http://hal.inria.fr/inria-00614523

[16] A. Wimmer, T. Jungel, M. Glück, and K. Dietmayer, "Automatic generation of a highly accurate map for driver assistance systems in road construction sites," in *Intelligent Vehicles Symposium (IV), 2010 IEEE*, June, pp. 281–286.

[17] A. Wimmer, T. Weiss, F. Flogel, and K. Dietmayer, "Automatic detection and classification of safety barriers in road construction sites using a laser scanner," in *Intelligent Vehicles Symposium, 2009 IEEE*, june 2009, pp. 578 –583.

[18] R. W. G. Hunt and M. R. Pointer, *Measuring colour*. John Wiley & Sons Sons, 2011.

[19] C. L. Franklin, *Colour and colour theories*. Arno Press (NY), 1973.

[20] L. M. Hurvich and D. Jameson, "An opponent-process theory of color vision," *Psychological Review*, vol. 64, pp. 384 – 404, 1957.

[21] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust features (surf)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346 – 359, 2008.

[22] D. Lowe, "Object recognition from local scale-invariant features," in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 2, pp. 1150–1157 vol.2.

[23] L. Fei-Fei and P. Perona, "A bayesian hierarchical model for learning natural scene categories," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2. IEEE, 2005, pp. 524–531.

[24] S. Westland, C. Ripamonti, and V. Cheung, *Computational Colour Science Using MATLAB*. Wiley, 2012.

[25] J. Van Gemert, C. Veenman, A. W. M. Smeulders, and J.-M. Geusebroek, "Visual word ambiguity," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, pp. 1271–1283, July.

[26] H. Jegou, M. Douze, C. Schmid, and P. Perez, "Aggregating local descriptors into a compact image representation," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, June.

[27] C. Seiffert, T. Khoshgoftaar, J. Van Hulse, and A. Napolitano, "Rusboost: A hybrid approach to alleviating class imbalance," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 40, no. 1, pp. 185–197, Jan.

[28] Y. Freund and R. E. Schapire, "A desicion-theoretic generalization of on-line learning and an application to boosting," pp. 23–37, 1995.

[29] "The geograph project," August 2012. [Online]. Available: http://www.geograph.org.uk/

[30] A. Torralba, K. Murphy, and W. Freeman, "Sharing features: efficient boosting procedures for multiclass object detection," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 2, june-2 july 2004, pp. II–762 – II–769 Vol.2.

[31] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: A statistical view of boosting," 2000.