Image-based Vehicle Indexing for a Seaport Transportation Surveillance System

David A. Sadlier*, Paul Ferguson, Ciarán Ó Conaire, Noel E. O'Connor, Killian Doyle
CLARITY: Centre for Sensor Web Technologies, Dublin City University, Dublin 9, Ireland.
david.sadlier@dcu.ie*

Abstract

In this paper we describe and evaluate two methods underpinning a surveillance-based content management system, designed for monitoring and profiling freight-based vehicular traffic in a seaport environment. The 'InSPeCT' system captures video footage of passing vehicles, and uses tailored optical character recognition (OCR) to index the footage according to vehicle license-plates and freight codes. The system provides advanced search techniques for the efficient retrieval of records, where each captured vehicle is profiled according to captured imagery and its associated interpretations. Underpinning this system is a method for detecting the boundaries of individual transits within sustained traffic flow. Considering it desirable to extend the indexing functionality of the system beyond OCR, a colour-based vehicle indexing approach is proposed and evaluated. All evaluations take place in the context of a system deployed in a busy national seaport.

1. Introduction

Currently being deployed on a trial basis at a busy national seaport, the InSPeCT (Integrated Surveillance for Port Container Traffic) system is a profiling and data-management system that indexes vehicle traffic flow based on a number of complementary information sources. The target user-group is seaport authorities aiming to advance their transportation-intelligence capabilities towards enhanced risk assessment of port-bound vehicular traffic. The system aims to overcome the inefficiencies associated with traditional CCTV surveillance solutions, by indexing captured visual media, and profiling it according to machine-generated tags. Utilizing best-of-breed OCR-based solutions for automatic number plate recognition (ANPR) and automatic freight-code recognition (AFCR) [1], the system captures not only vehicle license-plate information, but also the freight-codes that uniquely identify inter-modal shipping containers that are routinely transported by seaport-bound vehicles. The system uses this information to tag the photographic evidence of the corresponding vehicles, making it retrievable on that basis. Given continued advances in both ANPR [3, 10, 14] and AFCR [15, 11, 8, 13] technology, the system is designed to be easily upgraded once the state-of-the-art advances in either of these fields.

Apart from the processes of data capture and interpretation, a critical component of the InSPeCT system relates to the task of detecting the boundaries of individual vehicle transits within sustained traffic flow, and correctly associating independently recorded items (i.e. vehicle and freight codes) that are logically paired. As the ANPR and AFCR processes are deployed to generate multiple data reads, owing to asynchronous camera capturing, coupled with non-uniform vehicle speeds, these can be difficult to accurately group together without leakage across the groups corresponding to neighbouring vehicle transits. Therefore, boundary detection (capture association) represents a fundamental challenge for the system, to which there is associated risk in terms of overall performance - hence the motivation towards evaluating its reliability, as reported in this paper. It is also desired to extend the functionality of the system towards supporting the indexing of vehicles beyond OCR-based analysis. To this end, a colour-based indexing scheme is proposed, which aims to automatically identify the representative colour of each vehicle captured. Incorporating this information would extend the modality of the system to the benefit of more flexible search, which exhibits particular relevance in a security-oriented context (e.g. when the searcher may only be armed with partial information, such as partial number plate and vehicle colour).

1.1. Related Work

The area of automatic vehicle colour recognition remains a challenging research topic. Beyond the recognized difficulties of working with outdoor imagery [6, 5], other challenges relate to complications attributed to the nature of vehicles, such as the proliferation of glass, reflective surfaces, and peripheral bodywork [6, 12]. Nonetheless, some encouraging results have recently been reported. For instance, Baek et al. [2] report considerable accuracy for their approach designed to label the vehicles according to five colour classes. It is based on normalized histogram rep-

resentations of HSV-parameterized vehicle images and the use of Support Vector Machines (SVMs). Chen et al. [5] propose a similar method based on a combination of SVMs and YCbCr histograms, using three classes. In [12], the authors discuss the benefits of improved object segmentation for vehicle colour recognition accuracy. Specifically, they propose an approach for disambiguating imperfectly segmented vehicle chassis from distractive artefacts (windshield, tires, etc.). Their analysis is based on assumptions about vehicle shape, and they show the benefit of accurate segmentation in terms of performance accuracy for a seven colour classification problem. Dule et al. [6] also describe a seven-dimensional colour class problem, but address it from the perspective of identifying the most favourable combination of data parameterization and classification technique, demonstrating the strength of Artificial Neural Networks combined with complementary colour spaces.

Although there are some impressive results presented, a drawback relates to the narrow range of vehicle-types tested. This is reflected in the license-plate based region-of-interest selection rules used in [6], the car-shape based segmentation paradigm proposed in [12], and the various illustrative examples provided in [2]. It remains unclear as to whether the approaches would perform equally well for vehicle-types other than (e.g.) cars, such as those likely to be associated with a seaport environment. For instance, larger vehicle-types such as HGVs may present challenges to these solutions due to their shape differences and their peripheral bodywork (metal grill-fronts, bumper-bars, etc). Whilst Chen et al. [5] do allude to a diversity of vehicle types within their dataset (i.e. car, van, HGV), the limited scope of experiments and the accuracy obtained suggests that further experimentation in this are remains valuable.

After describing the InSPeCT capture station in Section 2, Section 3 presents a discussion on the technique underpinning the traffic-flow segmentation component of the InSPeCT system, followed by an evaluation of same. Section 4 proposes an approach for vehicle colour classification, and describes its accuracy applied to a real-world data set captured by the InSPeCT system. Section 5 provides discussion on the results obtained.

2. Capture System Overview

The InSPeCT capture station uses a seven-point camera network for the purposes of monitoring a single lane of traffic.Specifically, it consists of the following (see Figure 1):

- 2 × ANPR cameras: WVGA infrared spectrum cameras with built-in infrared illuminators, dedicated to capturing images of front-mounted license plates;
- 2 × AFCR cameras: WVGA visible spectrum cameras with built-in 'white-light' illuminators, dedicated to capturing rear-mounted freight-code information;

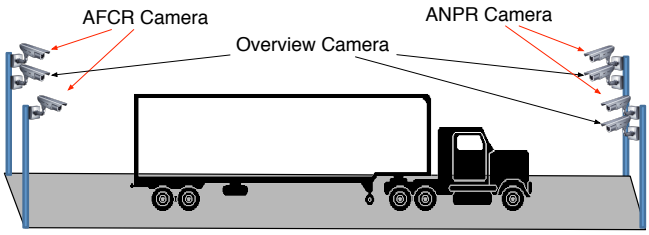

Figure 1. InSPeCT capture station.

- 3 × general purpose visible-spectrum cameras (termed 'overview' cameras) with built-in infrared for night-time operation targeted at capturing multi-purpose visible spectrum images.

The ANPR and AFCR cameras provide multiple data reads for each vehicle transit, and whilst this ultimately provides added robustness in terms of read accuracy, it results in multiple, possibly noisy incorrect reads that must be first associated with each other and corrected – this is the focus of Section 3. The overview cameras provide a useful source of data for vehicle analysis, such as the colour classification scheme described in Section 4.

3. Traffic-flow Segmentation

3.1. Overview

A key feature of the InSPeCT system is that it provides for non-invasive profiling, i.e. it is designed to provide for vehicle indexing in a free-flow environment, without requiring that vehicles need stop in order to be correctly captured. As mentioned in Section 1, a key task is therefore segmenting the traffic flow into a series of contiguous *events* – where an *event* is defined as a single vehicle transit (with/without freight) passing the system's camera installation, triggering the capture of associated imagery. The challenge of event *segmentation* is to automatically detect the boundaries of each individual event. To this end, the system analyses and combines the multiple ANPR/AFCR reads that occur as a consequence of a vehicle transit, to generate a redundancy-removed representation corresponding to the finalized event information (i.e. license-plate and freight code). Issues that need to be addressed are (i) the asynchronous timing of individual camera captures as the transit occurs, (ii) imperfections in the various OCR readings, and (iii) varying vehicle speeds.

3.2. Event Segmentation

3.2.1 Multi-capture grouping

In contrast to traditional OCR-based vehicle indexing systems that typically rely on a single capture from an OCR-camera, the InSPeCT system captures and interprets multi-

	0	1	D	1	2	3	7
0	88.2	0	0	0	0	0	0
2	0	0	0	0	86.3	0	0
D	0	0	88.9	0	0	0	0
1	0	97.6	0	91.3	0	0	0
2	0	0	0	0	77.5	0	0
2	0	0	0	0	90.0	0	0

Figure 2. Confidence matrix C for sample license plates **01D1237** and **02D122**

ple OCR reads and aggregates these towards a final result. This multi-capture approach has the benefit that if an individual capture happens to be incorrect, owing to the existence of additional captures, the system can still produce an accurate final read via the grouping process. Both the license-plate and freight-code captures are grouped on the basis of similarity measurement. Underpinning this is an algorithm for computing the similarity between two arbitrary-length character strings, which is based on exploiting the positional relationship of common characters and their corresponding confidence values. To illustrate, consider two license-plate reads of length M and N. We create a $M \times N$ confidence matrix, C, in which elements corresponding to identical characters are assigned the mean confidence value of the respective characters, and those of non-identical characters are assigned the value zero – see Figure 2. Dynamic programming is then applied to C as in Algorithm 1 to obtain the final similarity Sim (where $0.0 \leq Sim \leq 1.0$).

Algorithm 1 Two-capture similarity generator

procedure SIMILARITY($C[M][N]$)
  int a,b,i
  float Sim, $S[M][N]$
  $S = C$ ▷ S: score matrix
  for $b = (N-1)$ **to** 1 **do**
    for $a = (M-1)$ **to** 1 **do**
      $S[a-1,b] = \max(S[a-1,b], S[a,b])$
      for $i = 0$ **to** $(b-1)$ **do**
        $S[a-1,i] = \max(S[a-1],i], S[a,b]$
        $+C[a-1][b-1])$
      end for
    end for
  end for
  $Sim = S[0,0]/\max(M,N)$
  return Sim
end procedure

3.2.2 Grouping across cameras

For each OCR-camera (ANPR or AFCR), a 'queue' of capture groups is maintained, where each capture group corresponds to the set of all captures that are deemed to have originated from one particular vehicle or freight item, based on thresholding the similarity Sim between successive captures (threshold used = 0.8). Then, exploiting knowledge of the various camera positions, plus the expected time delay(s)/overlap(s) between when a vehicle would typically come into view for each camera (thresholds specific to each installation), the capture groups are compared across cameras (again, based on similarity thresholding), and on this basis grouped into event-level units. Figure 3 conveys this process using some sample data corresponding to the capture of license-plate and freight-code information from a synchronized pair of ANPR cameras, and a synchronized pair of AFCR cameras, where both pairs of cameras are monitoring a common traffic lane. The diagram illustrates the process of grouping captures from different cameras that are deemed to correspond to a common vehicle or freight item. Secondly, it illustrates that, despite the fact that there may be errors in various OCR readings, the system can still produce an accurate final capture based on group member 'voting' (majority rule). Thirdly, it illustrates the procedure of coupling together logically linked vehicles and freight (which is based on heuristics involving the temporal overlap of ANPR-AFCR capture groups).

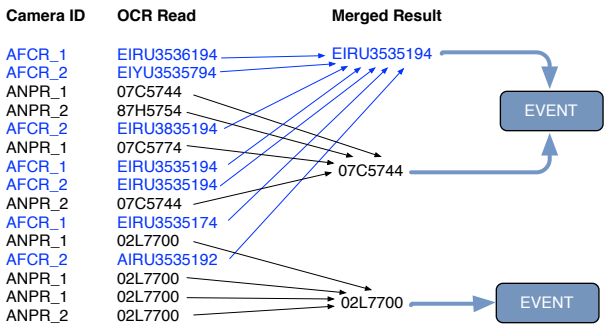

Figure 3. Grouping ANPR/AFCR readings from multiple cameras into distinct events.

3.3. Experiments and results

In order to evaluate the event segmentation system, a dataset was constructed corresponding to one-week's continuous capture (24/7) at the deployment site. The data exhibited a wide variety of lighting and weather conditions, as well as a variety of vehicle types. A manual ground truth was created for the correct segmentation boundaries and vehicle-freight associations, and the data comprised of over 1,200 vehicle events as a result. The results obtained are presented in Table 1. From this table, it is evident that for the case of ANPR, the system provided 98.5% Recall that a given vehicle event will be identified as distinct from a previous or subsequent vehicle event, (i.e. not erroneously

	Recall(%)	Precision(%)
ANPR	98.5%	99.9%
AFCR	100%	98.0%

Table 1. Results of event segmentation for ANPR and AFCR.

merged with a neighbouring event), with 99.9% Precision that an identified vehicle event will not be erroneously split into multiple sub-events. Corresponding values for AFCR (i.e. the segmentation of successive freight-items) were 100% Recall and 98% Precision.

For vehicle-freight association, it was calculated that, compared to the ground-truth, the system correctly associated 96.2% of all vehicle-freight pairs, only mis-associating the outstanding 3.8% in a number of particular circumstances – see Section 5.1.

Another important aspect corresponds to evaluating *event representation*. That is, assuming the accurate segmentation of an event, to what extent does the system attribute the correct license-plate/freight-code to the event, amongst the proliferation of correct/incorrect candidates (see Figure 3). Considering license-plates as a case in point, it turns out that out of the 1200+ events comprising the data corpus, 8.48% of these exhibited character errors in their corresponding OCR data. However, it was calculated that the grouping system handles over 95% of these cases.

4. Vehicle Colour Classification

As outlined in Section 1, vehicle colour classification represents a desired feature for the InSPeCT system. Clearly, it is not uncommon for some vehicles to exhibit more than one colour. However, for the purposes of this application, the dominant colour (that most frequently occurring) is considered representative. This section describes and evaluates a proposed approach for determining the dominant colour of vehicles captured in a day-time setting (natural illumination). Note, colour determination of night-time captured vehicles (artificially illuminated) represents future work.

4.1. Pre-processing

In order to improve the accuracy of subsequent colour classification, it makes sense to first attempt to localize the vehicle within the image, and then segment it from the background. From a practical point of view, whilst the variances in illumination and weather conditions associated with outdoor imagery tend to make accurate foreground segmentation difficult to attain, an approximation to an ideal result should nonetheless serve to increase the prospect of greater accuracy in subsequent colour analysis, owing to the exclusion of unrelated pixel regions.

(a) (b) (c)

Figure 4. (a) Background model [9]; (b) Car transit; (d) HGV transit

A series of methods were investigated; background modelling [9], mixture of Gaussians (MOG) [7], and clustering [4]. Based on some qualitative experimentation, it was found that for this stationary background scenario, aside from some minor result variances (e.g. the smoothness of illumination change handling), there was little difference in terms of the overall segmentation accuracy offered by each algorithm. Hence, owing to its superior computational efficiency, the background modelling method was utilized, whereby an up-to-date background image is maintained and is 'subtracted' from target images, revealing foreground objects. Note, the MOG approach might be preferable in situations where the background is at best quasi-stationary.

It was decided the background modelling would take the form of the selective running-average paradigm, which, helps prevent the background image being polluted by pixels that logically do not belong the background scene [9]. Furthermore, it was determined (based on some additional qualitative analysis) that for this particular scenario, better performance accuracy was achieved via aggressive background updating ($\alpha = 0.15$ [9]). Of course, the update parameter (α) could be tuned to reflect the characteristics (e.g. vehicle speed, etc.) of each particular traffic-flow scenario. Finally, a simple morphological filtering approach was used to remove noise (erosion followed by dilation - using 3x3 pixel kernels). Example segmentations are provided in Figure 4 for both a car and a heavy goods vehicle (HGV).

4.2. Classification

The method investigated for automatic dominant colour classification was based on a supervised machine learning approach. Specifically, a multi-class Support Vector Machine (SVM) was used, where each individual SVM was trained to discern a particular colour amongst the assortment of possible classes, in a one-versus-all manner. In practice, the colour corresponding to the SVM returning the highest confidence valued positive decision was then designated as the determined colour. Feeding this system was the pixel data constituting the vehicle images, processed into histogram format. Various approaches to the latter were investigated, but most favourable seemed to be HSV parameterized pixels, aggregated into 1-D histograms of hue (64 bins), saturation (16 bins), and value (16 bins).

Colour	Recall(%)	Precision(%)
Blue	75%	100%
Red	73%	100%
Black	95%	100%
Grey	85%	100%
White	80%	100%
AVERAGE	82%	100%

Table 2. Recall accuracy per colour for 100% Precision.

4.3. Experiments and results

For data collection, the InSPeCT system was activated for capture during the daylight hours over a period of one week. This resulted in the capture of visual images depicting 952 distinct vehicles. Upon examination of the data, it was observed that it was dominated by a mixture of vehicle-types exhibiting the following colours (in relatively equal proportions): red, blue, black, grey, white. Note, whilst yellow and green vehicles were also present, there was significantly fewer of these.

The experiments carried out were targeted at indicating the method's ability to distinguish between the prevailing colours observable in the dataset. The dataset was evenly divided between training and test data, yielding distinct data samples for both classifier training and classifier testing. Following the pre-processing of the images according to the method described in Section 4.1, the multi-class SVM system was then trained on the training data in accordance with manually annotated colour labels. The derived models were then evaluated using the (unlabelled) data points constituting the test data. The optimal SVM parameters were determined to be RBF kernel, with gamma=0.5, and cost=6. The results of this analysis are presented in Table 2, in the form of Recall accuracy per colour, rated at 100% Precision.

From the results, there is a clear disparity of performance across the five colours. For example, it is evident that the system had most difficulty in discerning red vehicles from the others (73% Recall), whilst it seemed to perform reasonably well in distinguishing black vehicles (95% Recall). Overall, the mean Recall performance (82%) would suggest that whilst encouraging, the approach, like others, falls short of closing the gap towards near-ideal classification.

5. Discussion

5.1. Traffic-flow segmentation

The evaluation of the traffic-flow/event segmentation process outlined in Section 3 indicates satisfactory accuracy for this component of the system. Particularly encouraging was the verified robustness of the approach in dealing with data captured under varying lighting conditions, from different types of vehicles, within a free-flowing traffic environment. Furthermore, the ability of the system to correctly combine multiple captures has proven to provide an effective means by which to correct instances of OCR errors, if and when they occur.

Nonetheless, despite the high performance observed so far, it may still be possible to further increase the accuracy by exploiting complementary data sources. That is, when the approach did fail, it was found to be largely attributed to a small number of vehicles performing unusual activity such as "tail-gating" (i.e. where a vehicle closely follows another vehicle to the extent that the licence plate is obscured). This type of behaviour could be recognized with the use of a hardware trigger, or, less expensively, via software-based visual analysis of the scene. Hence, this latter objective represents a future work target, the outcome of which should lend itself towards offering event segmentation performance that closes the gap towards 100% accuracy.

5.2. Vehicle colour classification

Deeper analysis of colour classification performances indicates that poor performance can be attributed to (i) the varying illumination conditions associated with outdoor capture, (ii) the complications associated with reflective vehicle surfaces, (iii) the proliferation of glass constituting the objects (windscreen and side windows), and (iv) the segmentation quality provided by the pre-processing step.

Adverse weather conditions present an unavoidable complication for image processing algorithms dealing with outdoor-captured imagery. In fact, manual examination of the data revealed significant instances of disparity in the hue for vehicles that exhibit the same 'colour'. In addition, the reflective surfaces presented additional complications for the system, via their tendency to mirror environmental objects and weather phenomena, especially during bright sunny conditions. Furthermore, the glass in the vehicles tends to present as a (grey-type) dark colour, which presents a further distraction. Finally, it was clear that object segmentation quality also has an influence over accuracy due to contamination from background and/or shadowing. Ideally, the vehicle would be segmented perfectly from the background environment, which would lend itself to more accurate colour classification, but in practice this ideal result can only be approximated – unless, as discussed in Section 1.1, assumptions can be made about vehicle-type [12]. Overall, whilst it has been shown that a reasonable level of accuracy for outdoor object colour classification can be achieved in a challenging environment with the method outlined, for the reasons discussed, closing the gap towards ideal performance remains a challenging task.

Future work will concern targeting an improvement in the results obtained. For instance, it may prove useful to attempt to determine the *nature* of the vehicle's colour ahead of classifying it. That is, if an initial distinction between

grey-type colours (i.e. white, black, and grey) and hue-type colours (e.g. red, blue, green, yellow, etc.) could be reliably made, it could be deployed as a pre-processing step towards improving the accuracy of subsequent identification (akin to eliminating non-relevant class possibilities from the next decision making process). To this end, preliminary experiments have already been initiated into methods for grey/hue-colour classification, and so far 80% accuracy has been achieved for the dataset, based on an approach using an SVM classifier fed with 1-D concatenated saturation-intensity histograms. Further experimentation may involve exploring other colour modelling techniques (e.g.) Gaussian Mixture Modelling.

6. Conclusions

In this paper, we described and evaluated two separate visual data analysis techniques, which underpin a visual surveillance content-management system for indexing and profiling vehicular traffic in a seaport environment.

The evaluation of the former, i.e. a traffic-flow segmentation technique that aims to associate multiple ANPR/AFCR reads with a single vehicle transit, revealed satisfactorily high accuracy, which represents a clear validation of the approach proposed. The second technique related to a proposed approach for vehicle colour classification, and whilst the evaluation of this method served to illustrate the challenges involved in bridging the gap towards near-ideal performance, the results obtained nonetheless represents a valuable benchmark for outdoor-captured, generic vehicle-type data.

Acknowledgements

The support of Enterprise Ireland under Innovation Partnership IP/2009/010 is gratefully acknowledged. This work is supported by Science Foundation Ireland under grant 07/CE/I114.

References

[1] ARH Inc.: Adaptive Recognition Hungary,. http://www.arhungary.hu/. 1

[2] N. Baek, S.-M. Park, K.-J. Kim, and S.-B. Park. Vehicle color classification based on the support vector machine method. In D.-S. Huang, L. Heutte, and M. Loog, editors, *Advanced Intelligent Computing Theories and Applications. With Aspects of Contemporary Intelligent Computing Techniques*, volume 2 of *Communications in Computer and Information Science*, pages 1133–1139. Springer Berlin Heidelberg, 2007. 1, 2

[3] S. L. Chang, L. S. Chen, Y. C. Chung, and S. W. Chen. Automatic license plate recognition. *IEEE Transactions on Intelligent Transportation Systems*, 5(1):42–53, 2004. 1

[4] T.-W. Chen, Y.-L. Chen, and S.-Y. Chien. Fast image segmentation based on k-means clustering with histograms in hsv color space. In *Multimedia Signal Processing, 2008 IEEE 10th Workshop on*, pages 322 –325, 2008. 4

[5] Z. Chen, N. Pears, M. Freeman, and J. Austin. Road vehicle classification using support vector machines. In *Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference on*, volume 4, pages 214 – 218, 2009. 1, 2

[6] E. Dule, M. Gökmen, and M. S. Beratoğlu. A convenient feature vector construction for vehicle color recognition. In *Proceedings of the 11th WSEAS international conference on nural networks and 11th WSEAS international conference on evolutionary computing and 11th WSEAS international conference on Fuzzy systems*, pages 250–255, Stevens Point, Wisconsin, USA, 2010. World Scientific and Engineering Academy and Society (WSEAS). 1, 2

[7] P. Kaewtrakulpong and R. Bowden. An improved adaptive background mixture model for real-time tracking with shadow detection. In *Proceedings of 2nd European Workshop on Advanced Video Based Surveillance Systems*, 2001. 4

[8] W. Pan, Y. Wang, and H. Yang. Robust container code recognition system. In *Proceedings of the 5th World Congress on Intelligent Control and Automation*, Hangzhou, China, June 2004. 1

[9] M. Piccardi. Background subtraction techniques: a review. In *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, volume 4, pages 3099 – 3104, 2004. 4

[10] M. T. Qadri and M. Asif. Automatic number plate recognition system for vehicle identification using optical character recognition. *Education Technology and Computer, International Conference on*, 0:335–338, 2009. 1

[11] Q. Tian and K. Wa. Container number recognition system. In *Innovation Magazine*, volume 1, 2006. 1

[12] Y.-T. Wu, J.-H. Kao, and M.-Y. Shih. A vehicle color classification method for video surveillance system concerning model-based background subtraction. In *Proceedings of the 11th Pacific Rim conference on Advances in multimedia information processing: Part I*, PCM'10, pages 369–380, Berlin, Heidelberg, 2010. Springer-Verlag. 1, 2, 5

[13] J. Yin, A. Cao, and J. Li. A new method for container code location. In *Proceedings of IEEE International Conference on Automation and Logistics*, Jinan, China, August 2006. 1

[14] L. Zheng and X. He. Number plate recognition based on support vector machines. In *Proceedings of the IEEE International Conference on Video and Signal Based Surveillance*, AVSS '06, Washington, DC, USA, 2006. IEEE Computer Society. 1

[15] H. ZhiWei, L. Jilin, M. Hongqing, and L. Peihong. A new automatic extraction method of container identity codes. In *Intelligent Transportation Systems, 2003. Proceedings. 2003 IEEE*, volume 2, pages 1688 – 1691 vol.2, 2003. 1