License Plate Recognition System by Using High Dimensional Model Representation

Nursinem Dere*[1], Burcu Tunga*[2]
*Istanbul Technical University
Department of Mathematical Engineering, Faculty of Science and Letters
Istanbul, Turkey
[1]derenu@itu.edu.tr
[2]tungab@itu.edu.tr

Abstract— **In this paper, it is proposed to match characters in content based images from real time scenes and extract the content as a text into the virtual environment by using the special algorithm called High Dimensional Model Representation (HDMR) for the system of License Plate Recognition. LPR is used to identify vehicles by reading license plates in image processing. Besides various techniques, a new matching algorithm is developed for the implementation of the LPR technology. LPR process is based on three major stages: Extraction of the license plate region from an image, segmentation of characters from the license plate region and recognition of characters which are segmented from the license plate. HDMR is used for working with high dimensional data based on image, at the stage of recognition of characters. The HDMR algorithm is used for matching the characters.**

Keywords— HDMR, license plate recognition, image processing, dimension reduction, character recognition

I. INTRODUCTION

License Plate Recognition (LPR) is known as a system that recognizes vehicles automatically. License plates are just like the identity of every vehicle on the road, they make vehicles recognizable and tell apart from each other which makes an each car unique. This feature provides many applications such as, traffic monitoring, applying traffic rules, parking spots, control of prohibited areas, and automatic toll passes [1, 2]. LPR system comprises of these stages: First stage is detecting license plate, then segmentation of license plate character and finally recognizing the character [3].

Extraction of the license plate region and segmentation of characters done by traditional extraction and segmentation methods. However, at the recognition of characters stage, unique implementation is used. The main reason of the recognition stage is to match and compare the characters of the plate region with the training set. The algorithm utilized in this paper to achieve great results from matching process is called HDMR. The success of the proposed algorithm is tested on vehicle images with different aspects and sizes.

HDMR is the method of dealing with high dimensional data based on images and disintegrate it into pieces to avoid complexity of calculations which arises from the fact of dimensionality. Moreover, it is used to decrease the error percentage and to make the software work faster on the process of recognizing and matching the characters of the license plate with the training set. HDMR decomposes a multivariate function which concludes constant variable, univariate functions and bivariate functions and a finite number of higher variate functions.

This paper is organized as follows. The second section is about the method we used in this paper, detailed explanation of HDMR with examples of how it is used. The main idea is explained at the third section. The third section is called Structure of LPR System and as its subsections Plate Region Extraction, Characters Segmentation and Character Recognition will be located in respectively. The fourth section consists of Experimental Analysis and Results of the LPR System. The final section covers the concluding remarks.

II. HDMR ON CONTENT BASED IMAGES

A. Mathematical Background

A multivariate function is divided into a number of less variate functions through the following HDMR expansion

$$f(x_1,\dots,x_N) = f_0 + \sum_{i=1}^{N} f_i(x_i) + \sum_{\substack{i,j=1 \\ i<j}}^{N} f_{ij}(x_i,x_j) + \cdots + f_{12\dots N}(x_1,\dots,x_N) \quad (1)$$

The right hand side HDMR components consist of an f_0, constant term, f_i, univariate terms, f_{ij}. bivariate terms and so on. To get each HDMR component, a different operator with multiple integrals are implemented to the both sides of the HDMR expansion given in (1) under a product type weight [4, 5]. The results are simplified through some normalization and vanishing conditions [4, 5]. Since this work deals with data instead of an analytical structure, the selection of the weight function becomes an important task. Dirac delta type weight is defined to obtain the HDMR components [4, 5].

978-1-5386-4184-2/18/$31.00 ©2018 IEEE

The formula of the first three HDMR components are derived as follows respectively

$$f_0 = \sum_{k_1=1}^{n_1}\sum_{k_2=1}^{n_2}\dots\sum_{k_N=1}^{n_N}\left(\prod_{i=1}^{N}\alpha_{k_i}^{(i)}\right) f\left(\varepsilon_1^{(k_1)},\dots,\varepsilon_N^{(k_N)}\right) \tag{2}$$

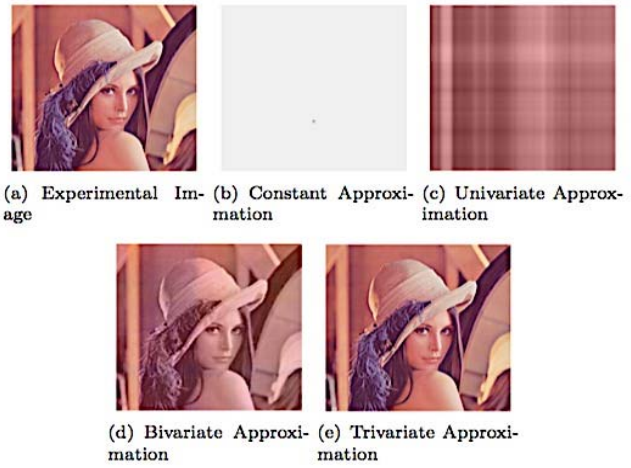
$$f_m\left(\varepsilon_m^{(km)}\right) = \sum_{k_1=1}^{n_1}\sum_{k_2=1}^{n_2}\dots\sum_{k_{m-1}=1}^{n_{m-1}}\sum_{k_{m+1}=1}^{n_{m+1}}\dots \sum_{k_N=1}^{n_N}\left(\prod_{i=1}^{N}\alpha_{k_i}^{(i)}\right) f(\varepsilon_1^{(k_1)},\dots,\varepsilon_N^{(k_N)}) - f_0 \tag{3}$$

$$f_{m_1m_2}\left(\varepsilon_{m_1}^{(k_{m_1})}\varepsilon_{m_2}^{(k_{m_2})}\right) = \sum_{k_1=1}^{n_1}\sum_{k_2=1}^{n_2}\dots\sum_{k_{m_1-1}=1}^{n_{m_1-1}}\sum_{k_{m_1+1}=1}^{n_{m_1+1}}\dots\sum_{k_{m_2-1}=1}^{n_{m_2-1}}\sum_{k_{m_2+1}=1}^{n_{m_2+1}}\dots \dots\sum_{k_N=1}^{n_N}\left(\prod_{i=1}^{N}\alpha_{k_i}^{(i)}\right) f\left(\varepsilon_1^{(k_1)},\dots,\varepsilon_N^{(k_N)}\right) - f_{m_1}\left(\varepsilon_{m_1}^{(km_1)}\right) - f_{m_2}\left(\varepsilon_{m_2}^{(km_2)}\right) - f_0 \tag{4}$$

where $1\le k_{m_1}\le n_{m_1}$, $1\le k_{m_2}\le n_{m_2}$, and $1\le m_1,m_2\le N$. The first component is a constant, the second includes one independent variable while the last one three variables.

When the HDMR studies are examined, it is necessary to use at most bivariate or trivariate terms for image reconstruction. Figure 1 shows the performance of these HDMR components. This figure includes the original image and its HDMR approximations.

HDMR is also considered as the multiway array decomposition method. That is, an image is decomposed by HDMR to get a constant, vectors and matrices. The constant term represents the average impact of the data, univariate terms represent the independent impacts of variables and the other terms with two or more variables represent the co-operative impact of related independent variables on the resulted HDMR approximation. Structurally univariate terms



(a) Experimental Image (b) Constant Approximation (c) Univariate Approximation

(d) Bivariate Approximation (e) Trivariate Approximation

Figure 1: Approximations of HDMR Components

are vectors and other terms with more than two variables are matrices.

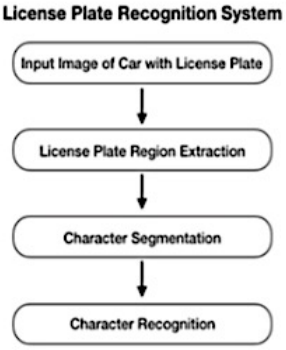


Figure 2: Stages of License Plate Recognition System

In this paper, we use the first two HDMR terms instead of original RGB image for dimensionality reduction. Using at most univariate terms is found sufficient because it has considerably low dimension and it is enough to represent the original image.

In this study, the new algorithm is developed by utilizing the HDMR method for the matching stage of LPR implementation. The details of the LPR system and the details of our new algorithm are introduced in the following section.

III. LPR USING HDMR

LPR is used to recognize a vehicle automatically. License plates are just like the identity of every vehicle on the road, they make vehicles recognizable and tell apart from each other which makes an each car unique. This feature provides many applications such as, traffic monitoring, applying traffic rules, parking spots, control of prohibited areas, and automatic toll passes [1, 2]. Figure 2 includes three basic steps. These can be considered as the detection of the plates, separation of the characters on the plates and the recognition of these characters.

The first task acquires a license plate containing image of a vehicle hence takes out the region of the plate. The second task separates the characters individually and the last task detects the individual characters and gives the result as the plate number. Here, we deal with character recognition part and we developed a new algorithm using univariate HDMR terms for the Turkish license plate characters. These plates includes 33 characters. 23 of them are Turkish letters and the rest 10 of them are numerals. These characters are stocked in our training set as matrices with the dimension of 42x24. The training set which is used in our algorithm is shown in Figure 3.

Figure 3: Training Set Characters

A. Plate Region Extraction

The proposed system uses some operations to find the regions includes characters in the input images and detects the most proper region as the location of the license plate.

1) Region Extraction Methods

License plate region extraction process is the first and the main step of LPR system. There are various approaches and methods for achieving license plate region extraction. Some of the different approaches for the stage of license plate region extraction can be exampled as Kim et. al. and E.R.Lee, P.K.Kim, H.J.Kim in which neural network based filters are used [6,7], S.K. Kim, D.W. Kim and H.J. Kim evaluated a segmentation which is based on genetic algorithm to get the license plate region [8], Park et. al. used the approach of vertical edge detection and vertical edge matching to extract Korean license plates [9]. Kahraman et al., and Tavsonoglu and Saatci, used Gabor filter in the detection stage [10, 11], Lotufo, Morgan and Johnson used OCR techniques [12]. Johnson and Bird managed to achieve the extracted license plate region by using template matching [13]. Choi utilized the approach based on vertical edge[14].

2) Stages of Extraction

The part of the extracting the region of license plate is the first stage of the whole system and it is important to get success in the next stage. The region has to be detected properly and completely for the next stages of the LPR system. After this process the license plate will be ready for the digit segmentation process.

Stages of extraction process which is used at this paper can be summarized as reading the image of the vehicle which includes the license plate of the vehicle, applying the methods of Sobel vertical masking to detect the edges and binarization, normalization, threshold based on Otsu Method, edge horizontal histogram [15], candidate plate regions detection by histogram, masking candidate plates, morphological operations [16], such as dilation to fill holes with vertical extension, dilation to fill holes with horizontal extension, and finally extracting location of the license plate after doing the binarization and smoothing processes.

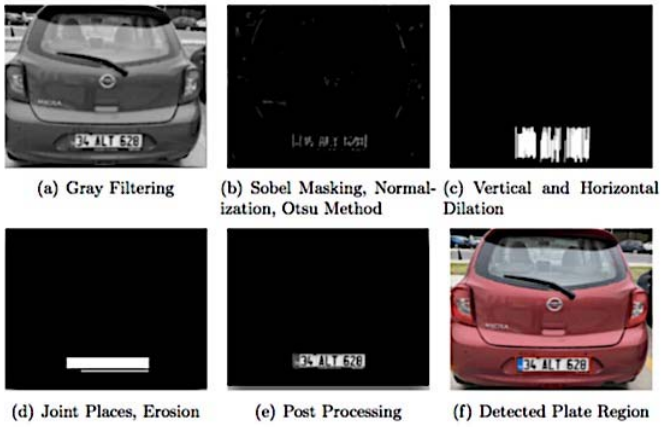
Right after attaining the extracted plate region successfully by the stages of the extraction, the LPR System continues for getting the characters segmented. There may be risks of incorrect detected regions of license plates.

The methods of the stages applying to the input image can be observed in Figure 4.

B. Character Segmentation

The second stage of the LPR System is character segmentation. To match the characters with the training set to get the characters of the plate, the regions of characters of the plate must be separated to be prepared for the third stage of this LPR System.

1) Character Segmentation Methods



(a) Gray Filtering (b) Sobel Masking, Normalization, Otsu Method (c) Vertical and Horizontal Dilation

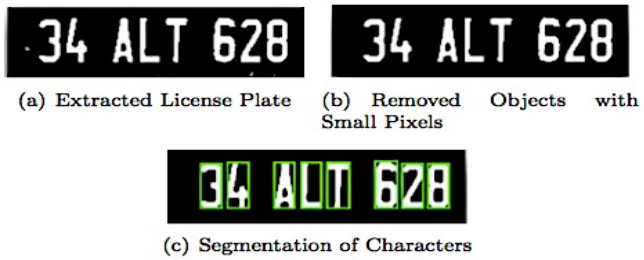
(d) Joint Places, Erosion (e) Post Processing (f) Detected Plate Region

Figure 4: Stages of License Plate Detection

There are several methods that have been developed and can be seen from the literature for the proposed stage of segmentation of characters such as blob colouring [1], projection histograms [17], partial differential equations (PDE) based technique [18], vertical margins [19], morphology, and connected component labelling (CCL) analysis [20]. All of these methods are served for the same purpose, to segment the characters to make them ready for the last step of the LPR system.

2) Stages of Segmentation

Blob detection algorithm is the algorithm selected to be used to get the segmented characters from the license plate. We apply binarization at the end of the extracting of the license plate stage once again to start the process. The blob detection algorithm can be considered as Connected Component Analysis (CCA) algorithm [15]. This algorithm indicates each of the characters by enclosing them in separate blocks starting from the left to the right. As the application of it can be seen at Figure 5, to avoid the small objects which are not representing any characters, we set a lower limit to consider the object that is detected in the segmentation is a character of a plate.



(a) Extracted License Plate (b) Removed Objects with Small Pixels

(c) Segmentation of Characters

Figure 5: Stages of Segmentation

C. Character Recognition

1) Character Recognition Methods

In the literature, for this stage, there are many techniques such as genetic algorithms [8], Markov processes [21], template matching [22], support vector machine [23], finite automata [24], artificial neural networks [8], [25] etc. are using. In this paper, HDMR method is used as a new perspective for the recognition stage of LPR system.

2) Stages of Character Recognition

The last stage of the License Plate Recognition System is character recognition. Amongst the other steps of LPR System, character recognition is the step used only in this paper by using the contribution of HDMR method.

(a) In Plate (b) In Training Set (c) In Matching Process

Figure 6: Types of Characters

In Figure 6, different types of characters dealing with in this project can be seen. Figure 6(a) represents character E from the image of license plate. Figure 6(b) represents binarized form of training character E that will be chosen as the reflection of the character 6(a) at the end of the character recognition stage. Figure 6(c) represents training character E at the perspective of the way of the matching algorithm.

Matching algorithm works between test character which is obtained after the segmentation of the license plate given in Figure 6(b) and the characters of training set given in Figure 3. These characters are 2-D binarized images. To apply the HDMR method on these images, first we have to add one dimension to the images [26]. Hence we obtain 3-D images that is, these images are considered as an array having 3 dimensions. When the HDMR method is applied to these 3-D images one constant (constant term), three vectors (univariate terms), and three matrices (bivariate terms) are obtained [26].

The size of the training characters and the segmented characters are fixed to 42x24. After adding dimension, sizes of each characters becomes 42x24x2. While calculating the univariate terms of HDMR, we obtained 3 vectors with the sizes of 1x42, 1x24 and 1x2 respectively. At the last step, we create a row vector by combining the vectors end to end. Hence, we obtain a new 1x68 row vector which identifies them all and is unique for each character.

The process of the preparation can be briefly narrated with a flow below:

1. For each segmented characters and characters of training set:

(a) Arrange the size of the character images as 42x24.

(b) Add third dimension to make the sizes of the binary images 42x24x2 [26].

(c) Calculate the univariate terms of HDMR by using the equation 3.

(d) Create a unique vector for each training and segmented characters by combining the univariate terms of HDMR to end to end.

2. Calculate the distance of the vectors of characters of the training set and vector of segmented test characters by using the Euclidean Distance Formula.

3. Determine the minimum distance amongst of the calculated distances, assign the training image as the value of segmented image.

4. Set the characters of license plate as a result.

The minimum value of the Euclidean Distance Formula gives the most approximated value of the comparison which means the closer matching among all others. This explains how the process ends at the stage of the Digit Recognition of License Plate Recognition System.

Table 1. Experimental Results

	Detection	Segmentation	Matching
Images	117/120	113/117	100/113
Accuracy	%97,5	%96,58	%88,49

IV. EXPERIMENTAL ANALYSIS AND RESULTS

To recognize Turkish license plate characters, the developed algorithm have been implemented to test the effectiveness of the LPR system. Vehicle images with the license plates which are used in the experiment have been taken in various illumination conditions, view angles and different image sizes.

Table 1. shows the results obtained from the algorithm developed for this experiment. In the experiment, among of the 120 vehicle images with license plates, 117 license plate regions of these are detected with the detection rate of %97,5. Meanwhile, three of them have spotted as a failure while detecting the license plate region.

The segmentation algorithm tested with 117 license plate images with 949 characters located by the previous step of the LPR. Among of the 117 detected license plates, we get 113 of them segmented successfully. The correctly segmented license plate character's percentage is %96,58.

All the images of the license plates of the vehicles are obtained from the license plates as black characters with white background.

Hence each character is obtained for recognition process using segmentation algorithm after that, HDMR method can be used for the character recognition. The recognition process using this method has been very successful. Of course, due to the extreme similarity of some characters (i.e. B and 8, 0 and D, A and 4, G and 6) very high success rates have not been achieved, but this increase has been left to further work.

To sum up, after with the %97,5 accurate extraction of the license plates, we are having %96,58 accuracy for the stage of segmentation before we apply matching. We see that successfully segmented plates are showing %88,49 success on reading all of the characters of the plate.

V. CONCLUSION

In this paper, it is proposed to get the license plate number from an image of a car by applying HDMR process. Considering the difference in sizes, backgrounds, camera angles, distances etc., the method we have developed has the correct matching with the rate of %88,49. The algorithm which is developed for this LPR application aims to recognize Turkish license plate characters. But yet, this system can be used internationally by adding new characters to the training set. The success rate of the system can be improved by enlarging the training set as well.

This project consists of 3 stages of LPR which are license plate region extraction, separating characters of license plate as character segmentation and character matching and we developed new algorithm for the third stage by using the method HDMR.

The system has been tested on a workstation featuring Intel Core i5 1.3 GHz CPU and the MATLAB R2018a is used.

REFERENCES

[1] S. Chang, L. Chen, Y. Chung, and S. Chen, *Automatic License Plate Recognition*, IEEE Transactions on Intelligent Transportation Systems (2012) 42-53.

[2] S. R. Aher, N. D. Kapale, *Automatic Number Plate Recognition System For Vehicle Identification Using Optical Character Recognition*, International Research Journal of Engineering and Technology (IRJET) (2017) 2515-2520.

[3] A. P. , *License Plate Character Recognition System using Neural Network*, International Journal of Computer Applications (2011) 36-39.

[4] M. Demiralp, High Dimensional Model Representation and its Application varieties, Mathematical Research, 9, (2003) 146-159.

[5] M. A. Tunga and M. Demiralp, *A new approach for data partitioning through high dimensional model representation*, Int. Journal of Computer Mathematics 12 (2008) 1779-1792.

[6] H. S. Kim et. Al., *Recognition of a Car Number Plate by a Neural Network*, Korea Information Science Society (1991) 259-262.

[7] E. R. Lee, P. K. Kim and H. J. Kim, *Automatic Recognition of a Car License Plate Using Color Image Processing*, International Conference on Image Processing (1994).

[8] S. K. Kim, D.W. Kim and H. J. Kim, *A Recognition of Vehicle License Plate Using a Genetic Algorithm Based Segmentation*, IEEE International Conference on Image Processing (1996) 661-664.

[9] Park et. Al., *Locating car license plates using neural network*, IEEE Electronics Letters (1999) 1475-1477.

[10] F. Kahraman, B. Kurt and M. Gokmen, License Plate Character Segmentation Based on Gobor Transform and Vector Quantization, ISCIS (2003) 381-388.

[11] V. Tavsanoglu, E. Saatci, *Feature extraction for character recognition using Gabor-type filters implemented by cellular neural networks*, IEEE International Workshop on Cellular Neural Networks and their Applications (2000) 63-8.

[12] R. A. Lotufo, A. D. Morgan and A. S. Johnson, *Automatic Number-Plate Recognition*, IEEE International Workshop on Electronic Image and Image Processing in Security and Forensic (1990) 4/1-4/8.

[13] A. S. Johnson, B. M. Bird, *Number-plate Matching for Automatic Vehicle Identification*, IEEE International Workshop on Electronic Image and Image Processing in Security and Forensic (1990) 4/1-4/8.

[14] H. J. Choi, *A Study on the Extraction and Recognition of a Car Number Plate by Image Processing*, Journal of the Korea Institute of Telematics and Electronics (1987) 309-3.

[15] L. Zheng, X. He, B. Samali, L. T. Yang, *An algorithm for accuracy enhancement of license plate recognition*, Journal of Computer and System Sciences (2013) 245-255.

[16] F. Faradji, A. H. Rezaie, M. Ziaratban, *A Morphological-based license plate location*, International Conference on Image Processing (2007) 57-60.

[17] H. Caner, H. S. Gecim, A. Z. Alkar, *Efficient Embedded Neural-Network-Based License Plate Recognition System*, IEEE Transactions on Vehicular Technology (2008) 2675-2683.

[18] J. Morel, S. Solemini, *Variational Methods in Image Segmentation*, Birkhauser, Boston.

[19] Y. Lu, *Machine printed character segmentation*, Pattern Recognition (1995) 67–80.

[20] K. Wang, T. Chia, Z. Chen and D. Lou, Parallel execution of a connected component labeling operation on a linear array architecture, J. Inf. Sci. Eng. (2003).

[21] Y.Cui and Q. Huang, *Character extraction of license plates from video*, IEEE Conf. Computer Vision and Pattern Recognition (1997) 502–507.

[22] M. Sarfraz, M. J. Ahmed, S. A. Ghazi, *Saudi Arabian License Plate Recognition System*, IEEE International Conference on Geometric Modeling and Graphics (2003).

[23] K. K. Kim, K. I. Kim, J. B. Kim and H. J. Kim, *Learning-based approach for license plate recognition,* IEEE Signal Processing Society Workshop (2000) 614–623.

[24] G. Adorni, F. Bergenti and S. Cagnoni, *Vehicle license plate recognition by means of cellular automata*, IEEE Int. Conf. Intelligent Vehicles (1998) 689–693. M. H. T. Brugge, J. H. Stevens, J. A. G. Nijhuis and L. Spaanenburg, *License plate recognition using DTCNNs,* IEEE Int. Workshop on Cellular Neural Networks and Their Applications (1998) 212–217. •

[25] M. H. T. Brugge, J. H. Stevens, J. A. G. Nijhuis and L. Spaanenburg, *License plate recognition using DTCNNs,* IEEE Int. Workshop on Cellular Neural Networks and Their Applications (1998) 212–217. •

[26] B. Tunga and A. Kocanaogullari, *Digital image decomposition and contrast enhancement using high-dimensional model representation*, Signal, Image and Video Processing (2018) 299–306. •

[27] MATLAB, *version 9.4.0 (R2018*a), The MathWorks Inc., 2018. •