Front. Comput. Sci., 2014, 8(4): 619–628
DOI 10.1007/s11704-014-3156-0

RESEARCH ARTICLE

# Road scene analysis for determination of road traffic density

**Omar AL-KADI (✉)[1], Osama AL-KADI[2], Rizik AL-SAYYED[1], Ja'far ALQATAWNA[1]**

1 King Abdullah II School for Information Technology, University of Jordan, Amman 11942, Jordan
2 College of Engineering and Computer Science, Australian National University, Canberra, ACT 0200, Australia



**Abstract** Road traffic density has always been a concern in large cities around the world, and many approaches were developed to assist in solving congestions related to slow traffic flow. This work proposes a congestion rate estimation approach that relies on real-time video scenes of road traffic, and was implemented and evaluated on eight different hotspots covering 33 different urban roads. The approach relies on road scene morphology for estimation of vehicles average speed along with measuring the overall video scenes randomness acting as a frame texture analysis indicator. Experimental results shows the feasibility of the proposed approach in reliably estimating traffic density and in providing an early warning to drivers on road conditions, thereby mitigating the negative effect of slow traffic flow on their daily lives.

**Keywords** road congestion, image texture, local binary pattern, scene morphology

## 1 Introduction

Regular and lengthy delays in traffic jams is not only a source of nuisance and frustration by wasting motorists' and passengers' time, but can also be a major cause for economic losses [1–3]. On the global level, the increase of air pollution and carbon dioxide by wasted fuel contributes to weather global warming. Down to the personal level, congestions may have a negative impact on motorists' health due to increased stress and road rage, late arrival which may result in disciplinary actions, passage obstruction of emergency vehicles, wear and tear on vehicle due to accelerating and breaking requiring more repairs, allocating more time for travelling which could be otherwise saved for other productive activities. All of these congestion collateral side-effects can be attributed to road capacity being below normal, free flow levels.

Traffic congestions could be attributed to many possible reasons. One of the most recurring circumstances is limited road capacity, especially at peak hours. Other reasons could be sudden increase in number of vehicles over a specific length in a road, accidents causing of blockage to lanes, roadwork that narrow the flow which result in bottlenecks in parts of the road, and bad weather conditions resulting in partial or complete roadway obstruction. This gives strong motivation for developing techniques that could reliably measure the healthiness state of real-time traffic flow, in analogy to blood pressure which gives indication of heart and circulatory system fitness.

Types of measurement technologies used for detecting road traffic density are twofold: intrusive and non-intrusive detectors. The former depends on signals emitted or reflected from passing vehicles, such as inductive loops, acoustic, magnetic, radar detection, while an example of the latter is video imaging detection which is categorised into tripwire systems that provide count and speed information at a single spot, and tracking systems relying on spatial information that can measure true density instead of detectors occupancy [4–6]. The main focus of this work is to improve video imaging tracking efficiency, as it has the advantage of providing spatial traffic information that are not related to a single point as in intrusive detectors.

Examples of recent major prior research related to real-time traffic analysis from video scenes for the purpose of congestion estimation include traffic conflict evaluation system

Received May 3, 2013; accepted February 20, 2014

E-mail: o.alkadi@ju.edu.jo

that was developed by Oh et al. [7] which looks for three categories in video images acquired from a single camera, which are traffic speed, trajectory, and conflict. In another work [8], Cheng and Hsu applied a time-varying adaptive system state transition matrix in Kalman filter for vehicle tracking and also utilised a regression analysis model for estimating traffic flow. Houben et al. extracted maximum phase congruency and edges from stereo images and matched together with local matching algorithm, then processed by maximum spanning tree clustering algorithm to group the points into vehicle objects [9]. Lane traffic characteristics, number of vehicles in each lane, and mandatory lane-changing fractions in lanes with traffic congestion were employed in this work. Bishop and Casabona used GPS-enabled networked mobile phones for collecting and analysing location based data set, and followed by sample filtering and route machining for traffic congestion determination [10]. Others like Ozkurt and Camci used a machine learning approach based on neural networks for traffic density estimation [11]. Based on maximum phase congruency and edges features matching and a 3D tracking, the system detects vehicles and determines their dimensions. Jain et al. applied a simulation-based analysis on a simple network topology where a local congestion protocol that controls the flow of traffic into near-congested regions for preventing congestion collapse sustaining time variant traffic bursts [12]. Chen et al. developed a night-time vehicle detection and traffic surveillance system, where vehicles head and tail-lights are located and processed by a spatial clustering for analysing the spatial and temporal features of vehicle light patterns [13]. While Yu and Chen used a simple consecutive temporal difference approach for traffic monitoring after pre-processing with a square neighbourhood algorithm for compensation of camera disturbances [14]. Also, Marfia and Roccetti proposed a short-term traffic congestion forecasting approach that requires no prior. knowledge on road condition, by estimating the time interval of how long would a congestion last [15]. Another work implements a vehicle counting method through blob analysis and tracking, and eventually determining the speed of vehicles [16]. A review on computer vision techniques utilised in traffic scene analysis can be found in [17].

Since image texture has shown in previous work its usefulness in differentiating between different image patterns [18,19], it could as well assist in representing a better understanding of road traffic scenes, if we take into account that the vehicles distribution are the patterns of interest. Therefore, and to the best of our knowledge, road traffic density from a texture perspective has not been classified before into advanced five stage conditions. It would be very advantageous from a computer vision system perspective to have a manifold view of road traffic density that provides road traffic control authorities with a clearer view on the traffic status quo, and hence giving time to take appropriate measures to deal with restricted flow conditions before it reaches to a complete stop. Also the system needs to be low-cost due to the large number of hotspots that would be monitored, and with minimal computational complexity to be able to work reliably in real-time.

In this work, a robust model that relies on both morphological analyses for estimation of average speed on roads along with video scenes randomness acting as an image texture measure was developed for measuring actual traffic flow. First step in the developed model is to pre-process each acquired video frame for non-uniform image intensity adjustment. Secondly, the vehicles are delineated and tracked by morphologically operating via motion segmentation, then estimating regional maxima, followed by connected component extraction, and this phase terminates by image opening and vehicle centroid determination. Thirdly, the average entropy of the local binary pattern (LBP) is used as a supporting measure for global measurement of the degree of randomness in the image texture; besides it is known for computational simplicity and robustness to monotonic grey-scale changes caused by illumination variation. Finally, the tracked vehicle speed is estimated and a simple multilevel thresholding classification approach is used for determination of road flow condition. The steps followed for the proposed model are illustrated in Fig. 1.

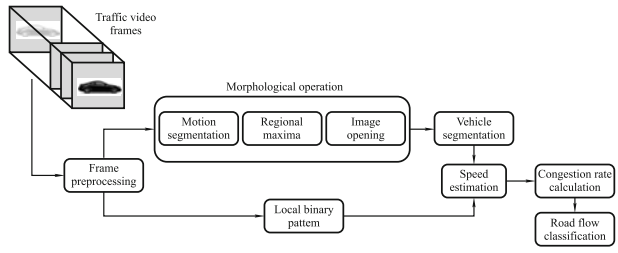


**Fig. 1** Road traffic density measurement model applied in this work

The paper is organised as follows. The cameras installation and video acquisitions are explained in Section 2, followed by the approach proposed for road congestion estimation in Section 3. The experimental results and associated discussion are presented in Sections 4 and 5, respectively. Conclusions are given in Section 6.

## 2 Materials

In order to ensure consistency and replication of results, the

camera type, specifications and installation procedure, and how the traffic videos were captured and made ready for subsequent processing are discussed next.

### 2.1 Camera specification and calibration parameters

The traffic video scenes were captured using a JVC Everio GZ-MG360 60 GB Hard Drive camcorder with 35x optical zoom (800x digital zoom) having a high-performance Konica Minolta lens. The camera was mounted on a tripod for stabilisation (i.e., reduction of camera shake) and appropriate elevation (i.e., better field of view) to achieve maximum sharpness.

At a height of 1.5 meters from the surface of a 5 meters high bridge, the camera tripod was fixed in the middle on one of the bridge sides overlooking the monitored road beneath. The camera was tilted by a 45° and focused down on the road to monitor a distance of 25 meters which is 15 meters away from the side of the bridge. Figure 2 illustrates the installation protocol followed in all traffic areas of study.

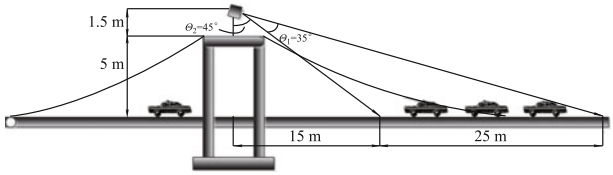


**Fig. 2** Camera installation for traffic video acquisition

Roads at designated study areas were composed of three lanes in each direction. Each video was further trimmed equally into two parts, one for incoming and another for the outgoing vehicles, giving us the opportunity to investigate the road condition for both directions from the same camera location without the need for a second camera.

The cameras were installed at carefully selected locations deemed as busy intersections in the city of Amman, Jordan. Monitoring the selected hotspots which are illustrated at the marked locations on the area study map in Fig. 3, can assist in the early diagnosis of traffic flow conditions, and thus can assist in studying, analysing and taking appropriate measures to alleviate congestions. The time of video capture corresponding to each hotspot area (A) covering the roads of interests (R) locations which could have multiple branches (a, b) are given in Table 1. This will facilitate results comparison in terms of road condition, given that rush hours in Jordanian cities range from 7:30 to 8:30 in the mornings and from 3:00 to 4:00 in the afternoons; off-peak otherwise.

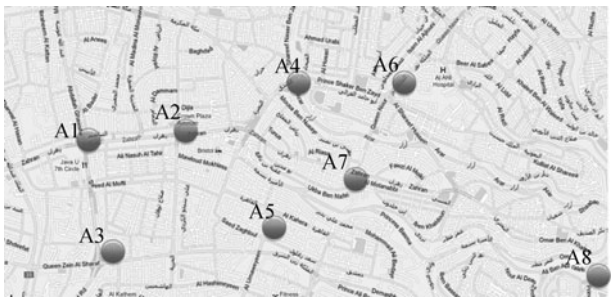


**Fig. 3** Camera locations covering hotspots at major busy intersection in Amman city, Jordan. Image courtesy of Google maps

**Table 1** Monitored roadway

| Road location | Capture time | Road location | Capture time | Road location | Capture time |
|---|---|---|---|---|---|
| A1-R1a | 16:00 | A2-R2a | 13:00 | A4-R3a | 14:00 |
| A1-R2a | 16:00 | A2-R2b | 13:00 | A4-R3b | 14:00 |
| A1-R2b | 16:30 | A3-R1a | 15:15 | A5-R1a | 08:00 |
| A1-R1b | 16:30 | A3-R1b | 15:15 | A5-R1b | 08:00 |
| A1-R3a | 15:45 | A3-R2a | 15:00 | A6-R1 | 12:30 |
| A1-R3b | 15:00 | A3-R2b | 15:00 | A7-R1a | 08:30 |
| A1-R4a | 15:00 | A3-R3 | 14:00 | A7-R1b | 08:30 |
| A1-R4b | 15:45 | A4-R1a | 13:30 | A7-R2a | 08:30 |
| A1-R5 | 15:45 | A4-R1b | 13:30 | A7-R2b | 08:30 |
| A2-R1a | 11:45 | A4-R2a | 14:00 | A8-R1a | 15:30 |
| A2-R1b | 11:45 | A4-R2b | 14:00 | A8-R1b | 15:30 |

### 2.2 Video acquisition and preparation

In order to preserve quality, the acquired video files were initially processed at full size with a resolution of RGB 720×480 pixels and 29 frames per second (fps), and the recording process was for duration of 3 minutes. This will result in a very large file size that would be deemed computationally intensive. Thus a preprocessing stage would be required to reduce the video size and frame rate while maintaining acceptable quality, a process that would have minimal side effect on the vehicle segmentation.

The captured videos from the fixed camcorder were in the MOD file extension, a JVC's implementation of MPEG-2 transport stream, and for the successive video frames to be appropriately read and processed, they were converted to AVI format, a common multimedia container used by different codecs and known for its good video quality and wide applicability. Xilisoft video converter http://www.xilisoft.com was used to convert the recorded MOD video files to its corresponding AVI format after setting appropriate resolution, frame rate and aspect ratio.

All acquired videos were converted to grey scale in order to reduce processing time since we are interested in detecting any car irrespective of its colour. Then each video frame rate and size was adjusted from 29 fps, 720×480 pixels having

a duration of 3 minutes to 15 fps, and 320×240 pixels and a 1 minute duration, respectively. The converted files had a size reduction of 80% with unnoticed loss in video quality, which resulted in a remarkable reduction in the video processing operation. Thereby, the videos are ready for the next step for vehicle detection and road condition analysis. The preprocessing and analysis program representing this work was written in MATLAB® Version 7.8 (R2009a) and tested on a 3 GHz Intel® Core™ 2 Duo processor notebook with 2 GB RAM memory running Windows® XP 32-bit operating system.

## 3 Methodology

This section starts by explaining the applied mathematical model for computing the average road velocity, then the video image vehicle detection approach along with the global LBP image average entropy is described, and eventually the road condition is classified on a scale of 1 (stopped flow) to 5 (free flow).

### 3.1 Representational model

Let $\Im$ be an order set of acquired video frames $v_i$, where $i$ is a certain point in the recorded time interval. We can define $C$ to be the set of univariate vectors $c$ that represents the number of detected vehicles in $\Im$, where $C = \{c_1, c_2, c_3, \dots\}$, $c_m \in \Im$. So the average road velocity $V_A$ could be estimated as:

$$V_A = \frac{D}{F_A}, \tag{1}$$

where $D$ is *a priori* known representing a fixed monitored distance that vehicles cross in meters, and $F_A$ as calculated in Eq. (2), is the average object flow in seconds required for the vehicles to cross $D$.

$$F_A = \frac{1}{n}\sum_{m=1}^{n}\frac{f_r}{C(c_m)}, \tag{2}$$

where $f_r$ is the frame rate and $C$ corresponds to the total number $n$ of $m$ detected vehicles $c$ per video frame.

For example in case A1-R1b in Table 1, suppose the monitored part of the lane $D$ was 25 meters long and the video displayed the frames at a rate of 15 per second for a duration of 1 minute and the number of detected objects (i.e., vehicles) was 740. The number of detected vehicles ($F_A$) in the video can be calculated by substituting in Eq. (2), so the average velocity ($V_A$) can be easily estimated by substituting in Eq. (1) which yields 74 km/hr.

Finally the state of the road denoted as congestion rate $CR$ can be estimated as:

$$CR = 1 - \frac{V_A}{V_{\max}}, \tag{3}$$

where $V_{\max}$ is the maximum road speed limit of 80 km/hr when substituting the value of 800 in Eq. (2), which denotes the number of detected vehicles, for a monitored distance of 25 meters for a duration of 1 minute.

### 3.2 Object recognition and scene analysis

#### 3.2.1 Image intensity adjustment

The video frames intensity values are adjusted by equalising the contrast of each image frame prior to the vehicle detection process. This pre-processing process is deemed necessary when capturing video in outdoors environments where varying illumination conditions are very common. The traffic videos should adapt to change in illumination during daylight and low-light, e.g., early mornings and late afternoons, therefore this process could assist in reduction of possible glare and shadows due to sunlight reflection and other changing weather conditions and in different seasons, which could affect image con trast, and hence the accuracy of vehicle segmentation.

A homomorphic filtering technique was employed for correction of the non-uniform illumination in the acquired video frames as defined in Eq. (4). It operates by assuming the scene $I(x, y)$ consists from an undesirable change in illumination $L$ due to varying lighting conditions at the time of video capture, and the scene reflectance $R$ related to physical properties of the objects, vehicles in our case. The illumination component $L$ is considering as additive noise, and hence can be filtered-out via a high-pass filter since it tends to have a gradual low-frequency change as compared to the abrupt high-frequency $R$ component.

The filter can be applied by performing an inverse Fourier transform to a high-pass filter $H(u, v)$ as $h(x, y) = \Im^{-1}[H(u, v)]$ and then convolving the image scene $I(x, y)$ in the log domain we get the intensity corrected image:

$$f(x, y) = \exp[h(x, y) \times \ln I(x, y)]. \tag{4}$$

#### 3.2.2 Morphological processing

• Spatial object motion segmentation

A simple, yet effective approach for detecting changes between successive video frames $f(x, y, t_1)$, $f(x, y, t_2)$, …, $f(x, y, t_n)$ is applying the absolute difference between each

frame and a reference frame $R(x,y)$, where $R(x,y) = f(x,y,t_r)$, and $r = 1$ denoting a zero state empty road traffic scene. Assuming that all video frames in the scene are appropriately registered and have the same size, spatial resolution and illumination conditions, and stationary objects would be eliminated and the moving image components would be preserved, described as:

$$D_{ij}(x,y) = \begin{cases} 1, & \text{if } |R(x,y) - f(x,y,t_k)| > T \\ 0, & \text{otherwise,} \end{cases} \tag{5}$$

where $k$ ($k > 1$) is the time of video frame capture, and $T$ is an appreciably different threshold set to the mid grey-level of the each video frame according to the following function:

$$T = \frac{1}{2}\left(\max(f(x,y,t_k)) + \min(f(x,y,t_k))\right). \tag{6}$$

• Regional maxima

Regional maxima can be defined as a certain connected components of pixels with a certain height attribute, such that all surrounding grey level pixels in the external boundary have strictly lower values. This could be utilised in removal of isolated low-valued points in the absolute difference image $D_{ij}(x,y)$. Isolated small structures are likely to be considered as noise, thus regarded as an essential procedure for reduction of false positives in the segmentation process. The elimination of such case can be done by arranging a certain set of $n$-connected regions (8 pixels was applied in this work) in $D_{ij}(x,y)$ and ignoring detected structures that are smaller in size. Nevertheless, small size and/or slow moving vehicles might be erroneously eliminated, yet the appropriate selection of the threshold value $T$ could contribute in mitigating this drawback, taking into consideration the camera field of view, depth of field, spatial resolution, focused distance, position and orientation.

• Connected component extraction

Connectivity is an important concept that defines regions and boundaries in images. To establish connectivity between two pixels in an image, they have to be neighbours and their corresponding grey levels satisfy a criterion of similarity.

Let $Y$ represent a connect component contained in a set $A$ and assume that a point of $Y$ is known. Then,

$$X_k = (X_{k-1} \oplus B) \cap A, \quad k = 1, 2, \ldots, n \tag{7}$$

yields all the elements of $Y$, where $X_0 = p$ ($p$ is a specific pixel in the image), $B$ is a suitable structuring element, and $\oplus$ is the dilation (XOR) operation.

• Image opening

Using a disk-shaped structure element $c$, which best resembles the shape of a car, with a radius $R$ set empirically to 4 pixels, a morphological opening on the binary image was performed for noise removal and other non-disk shape structures in the scene which are considered not related to vehicles of interest. The opening of each frame $f_i$, $i = 1, 2, \ldots$, by a structuring element $s$, denoted *fos* can be represented as [20]:

$$fos = (f \ominus s) \oplus s, \tag{8}$$

where $\ominus$ and $\oplus$ denote erosion and dilation, respectively.

• Image regions properties measurement

The area and centroid properties of each of the remaining objects in each frame are determined, and the object with the largest area is the detected vehicle.

Assuming a square $S_{QR}$, where $Q = (q_1, q_2)$ and $R = (r_1, r_2)$ would enclose each of the detected vehicles or objects $W_s$, where $s = 1, 2, \ldots, k$, such that $q_i \leqslant r_i$, where $i = 1, 2$, and the area of the object $A_s$ could be simply computed as $(r_1 - q_1) \times (r_2 - q_2)$. The centroid $CR_W(cr_1, cr_2)$ of $A_s$ can be defined as:

$$\begin{aligned} cr_1 &= \frac{1}{n}\sum_{i=1}^{n} W_{xi}, \\ cr_2 &= \frac{1}{n}\sum_{i=1}^{n} W_{yi}, \end{aligned} \tag{9}$$

where $n$ is the total number of pixels belonging to object $W_s$ within area $A_s$, $W_{xi}$ represents the $x$ coordinate of the $i$th pixel in $A_s$, $W_{yi}$ represents the $y$ coordinate of the $i$th pixel in $A_s$.

Finally the objects with the largest area would be considered the detected vehicles

$$W_{det} = \text{argmax}(A_{seg}). \tag{10}$$

This process when applied to a single video frame is illustrated in Fig. 4.

### 3.3 LBP

LBP is a local image texture descriptor that is robust to illumination variations, where it forms labels to image pixels by thresholding the neighbouring pixels with the centre value and converting the result to a binary number. This computational simplicity gives another advantage in making it suitable for real-time analysis.

For a local neighbourhood $N_{xy}$ of $P$ (where $P > 1$) number of pixels $g_{xy}$ in a monochrome image $I_{xy}$, a texture $T$ (where $T \in N_{xy}$) can be defined as a joint distribution of grey-level values of $g_{xy}$ [21,22]:

$$T \approx t(s(I_{x_0y_0} - I_{x_cy_c}), s(I_{x_1y_1} - I_{x_cy_c}), \ldots, s(I_{x_{P-1}y_{P-1}} - I_{x_cy_c})), \tag{11}$$


**Fig. 4** Example of vehicle detection and tracking in a video frame of road A1R5. (a) Acquired frame, (b) intensity adjustment, (c) background removal, (d) morphological processing and object detection, and (e) labeling detected vehicles on original frame for tracking subsequently speed measurement

where $I_{x_c y_c}$ corresponds to the grey value of the centre pixel of a local neighbourhood, and $I_{x_p y_p}$ $(p = 0, 1, \ldots, P-1)$ correspond to the grey values of $P$ equally spaced pixels on a circle of radius $R$ $(R > 0)$ that form a circularly symmetric set of neighbours. $s(x)$ is a threshold function that checks for the sign of the difference according as:

$$s(x) = \begin{cases} 1, & x \geqslant 0, \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

If the coordinates of $g_c$ where $(x_c, y_c)$ the coordinates of the neighbours $g_p$ are given by $\left(x_c + R\cos\left(\frac{2\pi p}{P}\right), y_c - R\sin\left(\frac{2\pi p}{P}\right)\right)$.

The spatial structure of the local image texture can be characterised by

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(I_{x_p y_p} - I_{x_c y_c}) \times 2^P. \tag{13}$$

In this work a $P$ and $R$ value of 8 and 1 was applied, respectively.

The average entropy measure $e_{\text{avg}}$ can be defined as:

$$e_{\text{avg}} = -\sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i), \tag{14}$$

where $z$ is a variable denoting LBP image grey levels and $p(z_i), i = 0, 1, \ldots, L-1$ is the corresponding histogram, where $L$ is the number of distinct grey levels. An example of a generated LBP texture image for a road scene is given in Fig. 5. Then the average LBP entropy ($e_{\text{avg}}$) is computed for each of the texture images and used for supporting the road flow condition estimation.

### 3.4 Classification by multilevel thresholding

A simple approach was employed to determine the road flow $RF$ based upon dividing the congestion rate on a scale 5 level scale, normalised from 0 to 1, by setting four different thresholds $T_s$, $(s = 0, 1, 2, 3)$ relevant to the road average velocity. The threshold describes how much the road capacity is utilised, where the congestion rate is represented as 1 minus the detected vehicles divided by the value of 800, which is the number of detected vehicles equivalent to the maximum road speed limit of 80 km/hr for a monitored distance of 25 meters for duration of 1 minute, which is indicated as:

**Fig. 5** A scene in road A1R5 and its corresponding average entropy LBP image

Let $T_s$ = [0.1, 0.3, 0.6, 0.8];
if $CR < T_3$ then RF = 5;
else if $CR \geqslant T_3$ AND CR $< T_2$ then RF = 4;
else if $CR \geqslant T_2$ AND CR $< T_1$ then RF = 3;
else if $CR \geqslant T_1$ AND CR $< T_0$ then RF = 2;
else if $CR \geqslant T_0$ then RF = 1;
end if. (15)

The feature vectors with the morphological operations are used to determine the road congestion rate can be summarised in Algorithm 1.

**Algorithm 1** Congestion rate estimation

**for** $i = 1, 2, \ldots, m$ **do** % Number of video scenes
$C = 0$; % Vehicle count initialisation
**for** $j = 1, 2, \ldots, n$ **do** % Number of video frames
$H_q = \exp(h_{xy} \times \ln(I_{xy}))$; % Perform homomorphic filtering
$D_{xy} = |H_q - I_r|$; % Absolute difference image
**if** $D_{xy} >$ S **then** $D_{xy} = 1$;
**else** $D_{xy} = 0$;
**end if**
$M = D$os % Perform morphological opening
**for** $k = 1, 2, \ldots, l$ **do** % Number of detected objects
$arg\max(M(k))$ % Find region maxima
**end**
$C = C + M$; % Compute number of detected vehicles
$F_A = f/C$; % Estimate average flow, where $f$ is frame rate
$V_a = D/F_A$; % Average speed, where D is monitored distance
$e_{\text{avg}} = H_{LBP} \times \log(H_{LBP})$; % Average LBP entropy, $H_{LBP}$ is histogram of LBP
$CR = 1 - (V_A/V_{\max})_{eavg}$; % Calculate congestion rate using video frame $e_{\text{avg}}$
$T_s = [0.1, 0.3, 0.6, 0.8]$; % Set threshold
**if** $CR < T_3$ **then** RF = 5; % Free flow
**else if** $CR \geqslant T_3$ AND CR $< T_2$ **then** RF = 4; % Moderate flow
**else if** $CR \geqslant T_2$ AND CR $< T_1$ **then** RF = 3; % Restricted flow
**else if** $CR \geqslant T_1$ AND CR $< T_0$ **then** RF = 2; % Slow flow
**else if** $CR \geqslant T_0$ **then** RF = 1; % Stopped flow
**end if**
**end for**
**end for**

## 4 Results

Running the system in eight different areas and for a total of 33 roads, Table 2 shows the classified road conditions after estimating congestion rates as discussed in the methodology section. Also the same table indicates the number of detect vehicles per each road, the detection error which is the percentage of error as compared to ground-truth, average road speed, and the LBP entropy which estimates the average randomness in the captured videos.

In order to make the interpretation of Table 2 easier and more visually appealing, the city roads traffic conditions were classified into five colours, as shown in Table 3, which determines the status of the roadway according to the proposed congestion rate estimation algorithm. The colour code would give the user a visual indication of the road status depending on the description of the road condition. The blue, green, yellow, orange, and red colours refer to the free, moderate, restricted, slow, and stopped flow conditions, while black colour code refers to the no-data condition, namely, either the road is clear where no vehicles are passing for a monitoring period of two minutes or a possible malfunction in the camera, the lens is covered or camera not operating.

The general daytime road conditions in the monitored ar-

**Table 2** Road condition estimation for 8 areas covering 33 different roads

| Monitored road | Detected vehicles | Detection error /% | LBP entropy | Estimated speed /km·hr$^{-1}$ | Congestion rate /% | Road condition |
|---|---|---|---|---|---|---|
| A1-R1a | 423 | 0.50 | 3.70 | 42.3 | 47 | Yellow |
| A1-R2a | 626 | 0.32 | 3.98 | 62.6 | 22 | Green |
| A1-R2b | 710 | 0.42 | 3.49 | 71.0 | 11 | Green |
| A1-R1b | 740 | 1.10 | 3.93 | 74.0 | 07 | Blue |
| A1-R3a | 81 | 5.19 | 4.79 | 08.1 | 90 | Red |
| A1-R3b | 449 | 0.22 | 4.12 | 44.9 | 44 | Yellow |
| A1-R4a | 738 | 0.00 | 4.21 | 73.8 | 08 | Blue |
| A1-R4b | 687 | 2.23 | 4.18 | 68.7 | 14 | Green |
| A1-R5 | 774 | 0.77 | 3.70 | 77.4 | 03 | Blue |
| A2-R1a | 214 | 3.38 | 4.48 | 21.4 | 73 | Orange |
| A2-R1b | 392 | 1.26 | 3.22 | 39.2 | 51 | Yellow |
| A2-R2a | 357 | 0.83 | 2.88 | 35.7 | 55 | Yellow |
| A2-R2b | 702 | 1.81 | 4.10 | 70.2 | 12 | Green |
| A3-R1a | 553 | 2.47 | 3.68 | 55.3 | 31 | Yellow |
| A3-R1b | 540 | 1.10 | 3.73 | 54.0 | 32 | Yellow |
| A3-R2a | 53 | 1.42 | 4.30 | 05.3 | 93 | Red |
| A3-R2b | 118 | 1.72 | 4.07 | 11.8 | 85 | Red |
| A3-R3 | 227 | 0.88 | 3.76 | 22.7 | 72 | Orange |
| A4-R1a | 433 | 0.92 | 3.00 | 43.3 | 46 | Yellow |
| A4-R1b | 541 | 1.99 | 3.17 | 54.1 | 32 | Yellow |
| A4-R2a | 279 | 0.36 | 3.74 | 27.9 | 65 | Orange |
| A4-R2b | 306 | 2.00 | 3.82 | 30.6 | 62 | Orange |
| A4-R3a | 192 | 2.67 | 3.80 | 19.2 | 76 | Orange |
| A4-R3b | 236 | 2.48 | 3.62 | 23.6 | 71 | Orange |
| A5-R1a | 690 | 0.00 | 3.40 | 69.0 | 14 | Green |
| A5-R1b | 357 | 3.48 | 4.03 | 35.7 | 55 | Yellow |
| A6-R1 | 210 | 0.00 | 4.31 | 21.0 | 74 | Orange |
| A7-R1a | 186 | 1.59 | 3.29 | 18.6 | 77 | Orange |
| A7-R1b | 663 | 2.21 | 3.76 | 66.3 | 17 | Green |
| A7-R2a | 597 | 1.97 | 3.27 | 59.7 | 25 | Green |
| A7-R2b | 538 | 0.00 | 3.27 | 53.8 | 33 | Yellow |
| A8-R1a | 735 | 1.94 | 4.14 | 73.5 | 08 | Blue |
| A8-R1b | 360 | 4.00 | 3.33 | 36.0 | 55 | Yellow |

**Table 3** Road traffic colour code interpretation

| Colour code | Description | Road congestion rate /% |
|---|---|---|
| Blue | Free flow | Less than 10 |
| Green | Moderate flow | Between 10 and 29 |
| Yellow | Restricted flow | Between 30 and 59 |
| Orange | Slow flow | Between 60 and 79 |
| Red | Stopped flow | Greater than or equal to 80 |
| Black | No data | Not available (0) |

eas are illustrated in Fig. 6, and Table 4 indicates the road conditions and corresponding peak status. Calculations related with daytime and rush hours, namely low number of detected objects in rush hours could mean congestion while in off-peak hours could mean empty or running road.

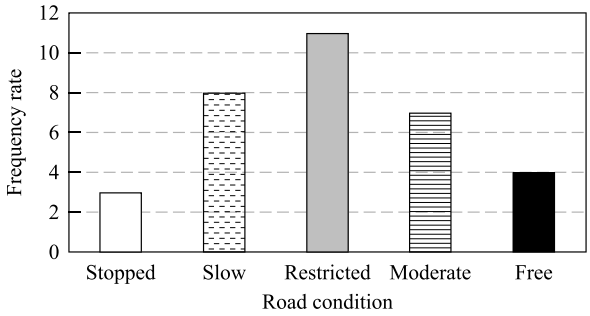


**Fig. 6** Road flow conditions for 33 roads in eight different busy areas in Amman city at different times recorded at daytime

**Table 4** Peak conditions for monitored roads in Table 2

| Colour code | Peak cases /% |
|---|---|
| Free | 50.00 |
| Moderate | 42.86 |
| Restricted | 54.54 |
| Slow | 85.71 |
| Stopped | 66.67 |

As a second measure that can assist in improving road condition estimation, the LBP entropy can measure the degree of randomness in the monitored roads without counting the number of vehicles in the assigned two minutes period. For instance, Figure 7 shows how the LBP average entropy can be used to differentiate between restricted and moderate flow road conditions for road A1-R1. Also when applied to six different roads in area A1 as illustrated in Fig. 8, it could assist in distinguishing between stopped, restricted, moderate and free flow road conditions.

The proposed method was also benched marked against

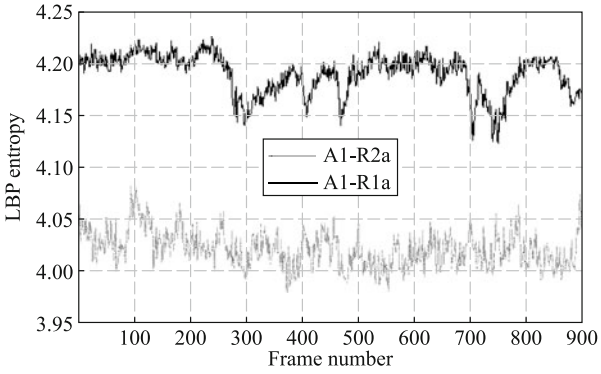


**Fig. 7** Congested LBP entropy A1-R1a (restricted flow ) vs normal LBP entropy A1-R2a (moderate flow)

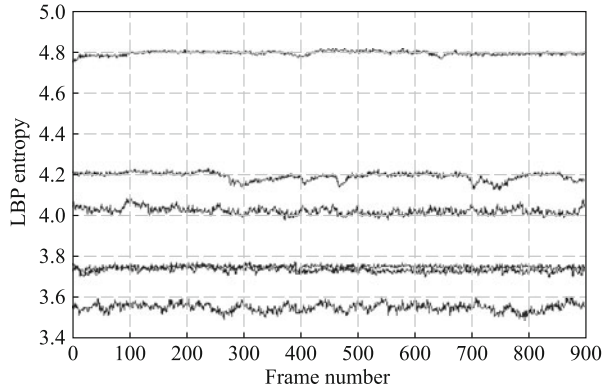


**Fig. 8** Time series LBP entropy for (from top to bottom) A1-R3a (stopped), A1-R1a (restricted), A1-R2a (moderate), A1-R2b (moderate), A1-R3b (restricted), and A1-R5 (free), respectively

two other statistical-based texture methods, namely grey-level co-occurrence matrix (GLCM) and run-length matrix (RLM), and with non-textural processed video frames for investigating robustness. The errors for all cases are shown in the boxplot of Fig. 9.

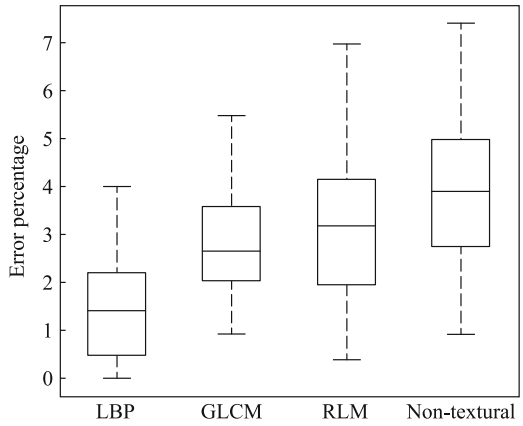


**Fig. 9** Boxplot of vehicle error detection for the proposed LBP method as compared to two other statistical-based texture method and non-textural processed video frames

## 5 Discussion

Road traffic usually exhibit stochastic behaviour due to conditions beyond our control, such as accidents, bad weather, and urgent road works that could affect the general traffic flow. Thus this work tends to provide a system that can automatically indicate the severity of road congestion, and hence improving the road flow by giving drivers an option to use an alternative way in case if the traffic situation is congested. As humans tend to be more attentive to colours, and are also mainly used in traffic lights to indicate road status and control traffic flow, this work utilises colour codes as well to further discriminate between different levels of congestion and to fa-

cilitate road condition's interpretation compared to numerical percentages.

It is expected that road conditions towards city centre in rush hours in early mornings to be moderate to slow flow, while the roads head out of the city centre to be free flow, and vice versa for the afternoon rush hours. Examining the bell-shaped diagram in Fig. 6, most of the road conditions were restricted having a percentage of 33.33% from the 33 monitored roads. While the slow and average flow conditions had a nearly similar percentage in the twenties, and 9.00% and 12.00% for the stopped and free flow, respectively. This indicates that most of the roads at the rush hours have a Gaussian distribution with restricted flow being the average condition, which is not unusual for that time of the day. Thus this system can assist drivers in avoiding stopped flow roads to an alternative road with a better condition, and hence mitigating the effect of road traffic congestion in Amman city.

There were a number of odd cases, such as A7-R2a and A7-R2b were classified as free flow although in peak time, this is because these roads are heading outside of the city centre. While road A2-R1a as slow flow although off-peak, which could be possibly due to road works. Table 4 shows also that the stopped, slow, and restricted flow conditions occurred mainly in the peak (rush hour) times that match when visually checking the captured road videos.

According to the proposed algorithm described in the methodology section, the classified road condition as stopped had always a high LBP value above 4, indicating a high randomness in the captured video, i.e., the presence of many vehicles in the road scene. However there were other roads that exhibited entropy values above 4, such as the cases A1-R4a, A1-R4b, A2-R1a, A2-R2b, A5-R1b, A6-R1, and A8-R1a when the road congestion condition was not stopped. This indicates that video scene randomness has to be integrated with the morphological road scene analysis for efficient road flow estimation. LBP average entropy is considered as a second measure for checking road conditions that relies on general image structure, and the more vehicles in the scene the more chaotic the image becomes and hence would give higher entropy, and vice versa. This can be applied also to distinguish stopped flow from empty roads, situations as in official holidays, late nights or very early mornings when usually roads are empty for long periods without cars passing. In both conditions the algorithm running on videos fed by monitoring camera assess no movements giving a stopped road condition, whereas the entropy can measure the degree of randomness in the video scene represented by the number of stationed vehicles superimposed on the empty road scene background. The edges of the vehicle structures would increase the randomness of the road background and hence give an indication of road occupancy. In other words, a higher entropy value would be noticed as compared when the road is empty. This is apparent from the comparison results in Fig. 9 where the LBP average entropy had an improved performance as compared to other statistical textural methods and to the non-textural processed video frames. This implies that the decrease in the vehicles detection error in respect to the other compared methods had a positive effect on the estimated average road speed, and hence the accuracy of the estimated congestion rate.

As there is no system free from disadvantages, the system currently does not indicate explicitly the reason for the increase in congestion (i.e., is it due to an accident, road work, double parking, etc.), and other traffic conditions and their possible effect on the approach proposed, such as bad weather, would be investigated in future work. Knowing the reasons behind congestion would give road traffic control authorities more information on the problem in order to take fast and suitable measures to deal with this situation. Also the level of road noise detected by the surveillance cameras microphones could be employed for improving congestion detection.

## 6 Conclusions

A new video scene analysis method was introduced for road traffic congestion estimation. The method combines both morphological operations for vehicle detection and speed estimation, and LBP entropy for video scene randomness measurement. The road traffic congestion is classified onto a 5-level scale that represents the various possible conditions of road flows. Results show the usefulness of this method when applied to different road flows at peak and off-peak times.

**Acknowledgements** The authors would like to thank Ruba Almumani, Basma Daaja, Rawan Abuosbaa, and Amani Madi for their assistance in acquiring the road traffic videos used in this work.

## References

1. Villagra J, Milanes V, Perez J, Godoy J. Smooth path and speed planning for an automated public transport vehicle. Robotics and Autonomous Systems, 2012, 60: 252–26
2. Towards a new culture for urban mobility. Green paper of European Commission, 2007, 1–6
3. Hartgen D, Fields M G, Moore A. Gridlock and growth: the effect of traffic congestion on regional economic performance. Reason Foundation, 2009, 1–8

4. Schrank D, Lomax T, Eisele B. 2011 Urban Mobility Report and Appendices. Texas Transportation Institute, 2011
5. Lomax T, Turner S, Shunk G, Levinson H S, Pratt R H, Bay P N, Douglas G B. Quantifying congestion. Transportation Research Board, 1997
6. Pattara-Atikom M, Pongpaibool P, Thajchayapong S. Estimating road traffic congestion using vehicle velocity. In: Proceedings of the 6th International Conference on ITS Telecommunications Proceedings. 2006, 1001–1004
7. Oh J, Min J, Kim M, Cho H. Development of an automatic traffic conflict detection system based an image tracking technology. Transportation Research Record, 2009, 45–54
8. Cheng H Y, Hsu S H. Intelligent highway traffic surveillance with self-diagnosis abilities. IEEE Transactions on Intelligent Transportation Systems, 2011, 12: 1462–1472
9. Houben Q, Diaz J C T, Warzee N, Debeir O, Czyz J, Insticc. Multi-feature stereo vision system for road traffic analysis. In: Proceedings of the 4th International Conference on Computer Vision Theory and Applications. 2009, 554–559
10. Bishop B, Casabona J. Automatic congestion detection and visualization using networked GPS unit data. In: Proceedings of the 47th Annual Southeast Regional Conference. 2009, 79
11. Ozkurt C, Camci F. Automatic traffic density estimation and vehicle classification for traffic surveillance systems using neural networks. Mathematical&Computational Applications, 2009, 14: 187–196
12. Jain V, Sharma A, Subramanian L. Road traffic congestion in the developing world. In: Proceedings of the 2nd ACM Symposium on Computing for Development. 2012, 11
13. Chen Y L, Wu B F, Huang H Y, Fan C J. A real-time vision system for nighttime vehicle detection and traffic surveillance. IEEE Transactions on Industrial Electronics, 2011, 58: 2030–2044
14. Yu Z, Chen Y P. A real-time motion detection algorithm for traffic monitoring systems based on consecutive temporal difference. In: Proceedings of the 7th Asian Control Conference. 2009, 1594–1599
15. Marfia G, Roccetti M. Vehicular congestion detection and short-term forecasting: a new model with results. IEEE Transactions on Vehicular Technology, 2011, 60: 2936–2948
16. Brahme Y B, Kulkarni P S. An implementation of moving object detection, tracking and counting objects for traffic surveillance system. In: Proceedings of the 2011 International Conference on Computational Intelligence and Communication Networks. 2011, 143–148
17. Buch N, Velastin S A, Orwell J. A review of computer vision techniques for the analysis of urban traffic. IEEE Transactions on Intelligent Transportation Systems, 2011, 12: 920–939
18. Al-Kadi O S. Combined statistical and model based texture features for improved image classification. In: Proceedings of the 4th International Conference on Advances in Medical, Signal and Information Processing. 2008, 1–4
19. Al-Kadi O S. Supervised texture segmentation: a comparative study. In: Proceedings of the 2011 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies. 2012, 1–5
20. Gonzales R C, Woods R E. Digital Image Processing. Prentice Hall, 2002, 299–300
21. Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24: 971–987
22. Guo Y, Zhao G, Pietik M. Discriminative features for texture description. Pattern Recognition, 2012, 45: 3834–3843