Multi-Camera Collaborative Network Experimental Study Design of Video Surveillance System for Violated Vehicles Identification

Hasan Thabit Rashid 1 and Israa Hadi Ali 2

1, 2 Department of Software, college of Information Technology, University of Babylon, Iraq
1 Computer Department, Education College for Women, University of Kufa, Iraq

E-mail: hasant.kurmasha@uokufa.edu.iq

Abstract: In this paper, we propose an experimental study of multi-camera collaborative network for surveillance the highway traffic turn in real life scenario, the target of surveillance based offline video processing is capturing the violated vehicles that driving violated paths in many cases specified by user-defined rules. Best topology of the experiment zone is considered and covered by four pillars; each has two (fixed and motorized) cameras that casing the entire specific effective field of view. As to author knowledge, there is no such available experiment, and hence, it could serve researchers that interested in. However, the experiment is done for around 180 recorded hours for 8 videos during 9 days, each video for one camera. It is designed based on the collaborative cameras principle for intelligent video surveillance systems and the outcomes show that the surveillance and the tracking of violated vehicles have been successes in most user-defined rules cases for more than 90% cases.

Keywords: Traffic violations; Area topology; Multi-camera collaborative network; Video analysis; Intelligent video surveillance systems.

1. Introduction

Nowadays, the cameras in use not only respond on the basis of the sensed data, but also support each other by their information, through multi-camera collaborative network MCCN. Coupling between two cameras may be affected geometrically and kinematic connections are the principal forms of coupling. To analyze low and higher resolution events, cooperation between static and dynamic cameras with their specific parameters is important. Field of view FOV is the maximum viewable angle from a camera. Spatial resolution is the ratio of the total number of pixels on its image element excited by the projection of an object in the real world. Higher spatial resolution may capture more data and makes pictures clearer. Depth of field is called the sum of distance between nearest and farthest objects that appear in an acceptably sharp focus of an image. To automatically extract valuable information from video data, intelligent video surveillance systems IVSSs is required with more scalability, flexibility and efficiency. IVSSs technology was used to monitor traffic, identify road collisions and control crowd activity in public spaces. The primary aim for surveillance is to retain understanding of the situation by fusing information from various cameras into a cohesive model of people, actions, and events to help a remote user understand anomalous and irregular activities [1, 2].
Single camera is weaker for tracking because it is covering a small area of such environment. It has problems such as segmentation of objects, occlusion of objects, etc. Multi-camera coordination had, until recently, allowed overlapping FOV. MCCN Setup works more efficiently with occlusion. Occlusion can be managed as the job of monitoring every single object by switched to another camera which can clearly realize it. Coordination and monitoring of MCCN performs many tasks; capturing and analyzing videos obtained from multiple cameras; Fuses the important information gained from different network cameras; it carries out the control actions necessary for the monitoring task in question [3-6]. However, we introduce short introduction at first section, and the traffic experimental considerations with weaknesses is arranged on second. The methodology and the topology of the experiment showed at the third section. The results and discussions, and conclusion have been explained consequently at last two sections.

2. Traffic experimental considerations

According to [7], many situations need to be addressed as to the surveilled area as following. The state transportation departments also maintain traffic cameras which simply track vehicle flows on roads and major arteries to conclude the valuable data. As well as monitoring the roads for collisions or significant delays, traffic camera footage is instrumental in future road growth and construction decisions. Monitor rules, cameras used to monitor speed and red light rules efficiently record driving violations and issue tickets. Promote safe driving, the visible intersection security cameras will encourage healthy driving practices and prevent moving violations. Weather-Traffic cameras are vulnerable to weather harm when they are tracking intersections or checking for traffic jams. The traffic security camera can damage, or destroy, heating, wind, rain, snow and ice. Accidents, there is also the risk that accidents will damage traffic cameras as they are placed on busy roads and intersections. Traffic surveillance cameras and red light or speed cameras have different functions and should therefore be regarded separately during set-up. For installing traffic control in right way, one may need to answer the following questions, what are your area's main roads? When is traffic the heaviest in your area (Rush Hour) at what time? Is traffic actually congested by other features on the roads? However, you need to install cameras to overlook specific areas of congestion and to make certain that all camera lanes are clearly clear and have a good view. Camera bodies can help protect the sensor from weather conditions by temperature and humidity. All above situations are considerations for specific surveillance systems which is very important to be taken into account when such system designs.

3. Methodology

3.1. Experiment design and considerations

Wired video system design is reliability and confidence; the ensuring to mix and match the most appropriate system components from different vendors as well as the power to optimize or expand the system to any size which demands quite knowledge. We need to know what the different components and how they work, and how they interact to construct a good reliable and perfect system. At the same time, the appropriate network and storage solutions depend greatly on the selected cameras, the camera settings such as resolution, compression, frame rate, and the number of cameras. One advantage of such wired system is to avoid the bandwidth and storage problems over complex and dynamic scenes on the applied scenario, so you may have to work on easy solutions, reducing the frame rate, lowering the resolution, or increasing the compression as to your proposed system requirements and best performance [8]. The considerations of our experiment include the following. We use outdoor-ready bullet cameras in general because of the environment of our scenario is noisy on traffic highway road. We use fixed cameras to keep all field of view FOV in watching, and (motorized, or PTZ) cameras with a large optical zoom to give us high-quality images for an effective zone in the work area especially for capturing the platelet numbers of vehicles, and to balance the cost of the budgets. Other considerations are involved within the devices such as day and night light-
sensitivity measurements, tamper-proof requirements, discreet installation, covering the area, FOV detection with high-details images for objects like license platelet recognition, analytic the system in general. To represent one meter of a scene, you need to 70-100 pixels in general, but for more details images such as human faces recognition or license platelets numbers' capturing and identification, operations like this needs to increase as many as 500 pixels per one meter. For license platelets numbers identification on vehicle with size 2×2 meters, we need to use one megapixel camera resolution or more. As to Fredrik Nilsson [8] in chapter (17), practices for indoor dome camera 2.8 mm lens mounted at 3 meters with straight FOV founded that the 125 pixels per meter is for face recognition while more than 250 pixel per meter is for face identification as to figure 1.

![Figure 1](image)

**Figure 1.** Person standing 6.5 meters away can be identified when the camera is mounted at a height of 3 meters. The horizontal coverage of the camera is shown in (a) and the vertical coverage in (b).

For positive identification of such small details of license platelets numbers or human faces and based on pixel density of horizontal dimension, Swedish National Laboratory of Forensic Science SKL recommend 80 pixels across a face for identification, while the international standard published by CELENCE, recommend a sufficient 40-100 pixels [9]. Because of most license platelets have the same width which is around 0.16 meter, then measuring its pixels will be easy as to the following figure 2.

![Figure 2](image)

**Figure 2.** The impact of resolution increasing over different pixels number.

For calculation, we need the surveillance area dimensions which are scene width SW, scene high SH, the number of pixels needed for platelet resolution PR, and platelet width PW. Now, we can compute the number of pixels for the scene resolution horizontally SRH and vertically SRV as to the following Equation (1) and (2).

\[
SR(H) = \frac{SW}{PW} \times PR \text{ (width)} \tag{1}
\]
\[ SR(V) = \frac{SH}{PH} \times PR(high) \]  

Now, to calculate SRH as to Equation (1), the general platelet dimensions are (20×10) (width×high) cm and assume that SW equal to (4 m × 125 pixel / 20 cm) × 80, the result is 2,000 pixels. For SRV calculation, by assume SW equal to (0.70 m × 250 pixel / 10 cm) × 40, the result is 700 pixels. In the scene, to identify the platelet number with high resolution based on above, we need 2000×700 pixels resolution which stands for 1080p to PTZ or Motorize cameras (optical zooming) while fixed 5MP 2560×1920 pixels (digital zooming) over the entire scene resolution to identification the vehicles. The general dimensions measures as standard for high definition HD videos are as following. 720p (1 MP) 1280×720 HD equal to 921,600 pixels (resolution = 922 of kilo pixels), 1080p (2MP) 1920×1080 Full HD (FHD) equal to 2.07 million pixels, 2048 × (unspecified) (2 K) equal to more than 2,000,000 pixels (cinema resolution), 1440p (4MP) 2560×1440 Quad HD (QHD) equal to 3.686,400 pixels, 2560×1920 (5MP) equal to 4,915,200 pixels, 2160p (4 K) (3480×2160) Ultra HD (UHD) equal to 7,516,800 pixels, and 4680p (8K) (7680×4320) as in figure 3[10]. However, by determining the required resolutions, the calculation of the number of cameras needed to cover the surveillance area can be identified. Number of cameras needed is equal to horizontal scene resolution divided by camera resolution. For our experiment, the vehicles size more than 2×1.5 meters and then 5MP camera will be more quite enough.

![Figure 3](image-url) 1080p (2MP) specifications.

3.2. Experiment topology

We test and select the best topology of cameras positions of our experimental system by four pillars, two cameras for each, four rules to detect the violated vehicles as in figure 4. The first rule is rule 1, vehicles that take the violated path (inverse direction) for cross over to the opposite side of the road straight from the area of pillar 1 to the best topology of cameras positions of our experimental system area of pillar 3 and vice versa. The second rule is rule 2, vehicles that take the violated path (inverse direction) for cross over to the opposite side of the road straight from the area of pillar 2 to the area of pillar 4 and vice versa. The third rule is rule 3, vehicles that take the violated path (inverse direction) for moving on the same side of the road straight from the area of pillar 2 to the area of pillar 1 and vice versa. The forth second rule is rule 4, vehicles that take the violated path (inverse direction) for moving on the same side of the road straight from the area of pillar 4 to the area of pillar 3 and vice versa.
Figure 4. Applied topology based the proposed user-defined rules (R1: rule one, R2: rule two, R3: rule three, and R4: rule four) of our scenario. Yellow rectangle: effective FOVs (optical motorized zone) watched by all four motorized cameras in the pillars.

For achieving that, we cooperate four fixed and four motorized cameras to capture the violated vehicles and its plate late numbers for identification purposes. Based on that, the vehicles tracking will be done based on birth and death zones and trajectories of them will be analyzed based on above rules to report each violated one. Each violated vehicle will be reported with its specific properties and metadata such as (time, numbers of frames, front side images of vehicle, rear side images of vehicle, platelets licenses images, etc.) using video processing and machine learning techniques such as Haar cascade classifier. Also, it is important to consider the relations of effective FOVs time and objects appearances between cameras over the scene cameras' motion, and scene structure for any applied tracking algorithm. Several relationships exist between the images with different views are very important for the construction of vehicles tracking. However, the performance of any IVSS is subject to the performance of its cameras, hardware costs, installation costs, maintenance costs and operating costs [11, 12, 13].

3.3. Experiment Setting up

The purpose of our experiment is to detect all violated vehicles such as car, bus, truck, and scooter that take a violated path to be turned on double inverse side of highway road. The detection will be done based on the conception of MCCNs on single device video recording DVR system via RG59 coaxial cable based on two stages. The first stage is for measuring the distances between each pillar that will carry up two cameras as in figure 5. Four pillars positioned on the rectangle figure, each carried up two cameras, the upper camera is fixed 5MP (at 3 meters of height) to covering the entire area while the lower one is a motorized or PTZ 2MP (at 2.75 meters of height) which set to capture platelets of vehicle. This scenario will help in covering the front and rear of the vehicle at same time. Capturing platelets of vehicles over specific recorded frames is done by setting the optical lens for the better length (4 – 5 meters) to be wider (catch misleading vehicles) and clear when recovering the
frames from both two opposite pillars. The fixed 5MP cameras are settled as to covering the same FOV for each pillar. DVR unit is used to control and simultaneous the time ticks and dates on frames of different recorded videos, and however, many positions and situations are tried and studied until the best setting up and shooting video is accomplished. However, the following Figure 5 shows all first stage representations.

![Figure 5](image.png)

**Figure 5.** First stage representations: area topology with distances and directions. Red arrows: pillars positions, blue box: control unit position.

The second stage is for DVR software installation and setting up different properties for all eight cameras where all viewed in one control monitor. For best recording process, we choose one hour time duration of each video (to be saved), 11 frame per second FPS for fixed cameras, and 13 (manufacturing setting) FPS for motorized or PTZ cameras. Table 1 below shows the setting options used in practices.

**Table 1.** The applied setting for cameras and video information.

| Camera                  | Frame rate | Width × Height | Video Type and Duration |
|-------------------------|------------|----------------|-------------------------|
| Fixed Cameras           | 11 FPS     | 1920×1080      |                         |
|                         | 11 FPS     | 2560×1920      | MPEG-4 / H.264 / 1 hour |
| Motorize or PTZ Cameras | 25 FPS     | 1920×1080      |                         |
|                         | 13 FPS     | 1920×1080      |                         |
4. Results and Discussions

We show how the recording videos are cooperated for the purpose of capturing the violated vehicles based on the proposed user-defined rule one (R1) and rule two (R2) with its platelets information identification, the times of violation, and other important metadata from all cameras as to their specific purposes into scenario and generating them as new reports. We present the cooperating benefits for avoiding the problems of overlapping effective E-FOVs such as full and partial occlusions, technology mistakes during recording time, and other situations.

4.1. Overall Scene

One violated vehicle based user-defined rule one (R1) of scene number 1926.jpg from all four fixed 5MP cameras (11 fps) and its corresponding 2276.jpg from all four motorized cameras (13 fps) on (29/09/2019 1:02:59 PM) as to the following figures 7, and 8.
Figure 7. Background scene over four fixed cameras at frame 1926.jpg. Top left: fixed camera installed on pillar 3. Top right: fixed camera installed on pillar 1. Bottom left: fixed camera installed on pillar 2. Bottom right: fixed camera installed on pillar 4.

Figure 8. Capturing platelets of violated vehicle (2276.jpg) that detected by R1. First row: two frames of PTZ cameras on optical motorized zone. Second row: two frames of motorised cameras on optical motorized zone. Third row: zoomed PTZ frames of first row. Note that platelets license form in Arabic language and numbers (50972 M / Baghdad, Iraq).
4.2. Occluded Scene

The following situations in figures 9 and 10 are for occluded scenes, the first one is harmed platelets case and the second one is for misleading vehicles case. It solved by other cooperated cameras on different positions.

**Figure 9.** Harmed Platelet of violated vehicle captured by rule two (R2) from one rear side only at frame (15412.jpg) in 29/09/2019 on 1:19:45 PM. Note that platelets license form in Arabic language and numbers (6406, Iraq, Basra).

![Image](image1)

**Figure 10.** Violated vehicle of inverse rule one (R1) from one front side only at frame 14802.jpg. Note that only one tire is present on PTZ motorized zone by pillar 3 camera while it is captured by PTZ camera of pillar 1 clearly. Date of recording is 2/10/2019 on 1:14:06 PM. Note that platelets license form in Arabic language and numbers (362081, Irbil, Iraq).

4.3. Technical and emerged mistakes

The founded noisy camera video output (factory errors) as in fixed camera of pillar 4 can be avoided by other cooperated cameras’ videos based MCCN principle and video processing techniques. The number of each frame for each camera should be equal but in the PTZ and motorized cameras and based on factory setting; they have 13 FPS while it is 11 FPS in fixed cameras. This problem is solved
as the calculation of the following Equations (3), (4), and (5) for fixed cameras and Equation (6) for motorize or PTZ cameras.

\[ \text{Current second (CS)} = \frac{\text{Frame Number(fixed)}}{11} \]  

(3)

\[ \text{First Frame at CS} = CS \times 11 \]  

(4)

\[ \text{Fixed Current Frame (FCF)} = \text{Frame Number} - \text{First Frame at CS} \]  

(5)

\[ \text{Motorize Current Frame (MCF)} = CS \times 13 + \text{FCF} \]  

(6)

As an example of above figures 8 and 9, the current second (1926÷11) equal to 175 seconds as to Equation (3) while the current frame in the current second is 1926 - (175×11) equal to 1 as to Equation (4) and Equation (5). The current frame in motorize camera is (175×13) +1 which equal to frame number 2276.jpg as to Equation (6). The fixed cameras on pillars 3 and 4 (both are 5MP) have noisy appearance video output (company errors) and they need to be processed, therefore the problem is solved by the opposite and cooperated cameras as to the same rules of violated. Over all the recorded videos of different daytimes, there are more than 90% cases are successes in capturing the violated vehicles among the user-defined rules. This will help effectively in vehicles detection and tracking using video processing and machine learning techniques. However, by solving and avoiding the technical and emerge errors of experiment setting; there will achieving 100 % success.

5. Conclusion

We propose a methodology of using techniques and technologies in cooperate with MCCN and area topology principles by explain how to take the important criteria into account for achieving a good IVSS results. Based on that, we design an experimental study following the mentioned scenario of traffic turn on highway in real life environment; the resulted recorded videos will be very helpful and serve in design and test IVSS over video processing and machine learning techniques. The topology of the experiment has been set as best as possible to eliminate the costs of financial, technical, and operational equipment's and hence; the four settled user-defined rules summarize the surveillance process of the violated vehicles by allowing to detect objects initially and tracking it over specific regions of interests in offline processing (to be identified) finally. However, the gathered raw data set will be available online for researchers to use and generate perfect IVSS for object detection, tracking, and identification.
References

[1] Kolekar, M. H. (2018). Intelligent video surveillance systems: an algorithmic approach. CRC Press. Chapter 7, pp 152-157.

[2] Andrew P. Tarko (2020), Techniques and technologies of observing traffic conflicts, Measuring Road Safety Using Surrogate Events, Elsevier, Chapter 4, Pages 47-65.

[3] Silva, T. C., Laiz, M. T., & Tabak, B. M. (2020). Traffic campaigns and overconfidence: an experimental approach. Accident Analysis & Prevention, 146, 105694.

[4] Zhang, J., Perloff, J. M., & Lu, F. (2020). Informing and inquiring: Experimental evidence on reducing traffic violations. Journal of Development Economics, 147, 102534.

[5] Landowski, M., & Landowska, A. (2019). Knowledge extraction from the experimental data of the vehicle traffic volume. Transportation Research Procedia, 39, 270-279.

[6] Natarajan, P., Atrey, P. K., & Kankanhalli, M. (2015). Multi-camera coordination and control in surveillance systems: A survey. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 11(4), 1-30.

[7] Website of videosurveillance (01/09/2020), Advantages of Traffic Surveillance Cameras, https://www.videosurveillance.com/traffic.asp.

[8] Nilsson, F. (2016). Intelligent network video: Understanding modern video surveillance systems. CRC Press. Chapter 17, pp 301-304.

[9] European Committee for Electrotechnical Standardization CENELEC Guides (10/08/2020), https://www.cenelec.eu/membersandexperts/referencematerial/cenelecguides.html.

[10] Website of Megapixel calculator (16/09/2020), https://toolstud.io/photo/megapixel.php.

[11] Hasan Th. & Israa H. A. (2019), Varying Techniques for Vehicles Violations over Intelligent Surveillance Systems: A Short Review. Jour of Adv Research in Dynamical & Control Systems, Vol. 11, 10-SP.

[12] Han, Z., Li, S., Cui, C., Song, H., Kong, Y., & Qin, F. (2019). Camera planning for area surveillance: A new method for coverage inference and optimization using location-based service data. Computers, Environment and Urban Systems, 78, 101396.

[13] Zahra Arbabi et al 2020 J. Phys.: Conf. Ser. 1530 012111.

Hasan Thabit Rashid Kurmasha is currently a lecturer with University of Kufa, Najaf, Iraq. He received his Bachelor degree in Computer Science from Babylon University, Hilla, Iraq in 2005, and his Master degree in Information Technology from UNITEN University, Malaysia in 2013. He is currently a PhD candidate in Babylon University. His research interests include Intelligent Video Surveillance Systems, Image and Video Analysing, Computer Vision.

Prof. Dr. Israa H. Ali received her PhD degree from University of Babylon in 2013. She is currently a Professor in Software Department, College of Information Technology, Babylon University, Iraq. Her main areas of research interest are Video Tracking and Analysis, Image Processing, Data Mining, Machine Learning, Multimedia and Information Hiding.