DESIGN FRAMEWORK FOR INTERNET OF THINGS BASED NEXT GENERATION VIDEO SURVEILLANCE

A Thesis Submitted to the College of Graduate and Postdoctoral Studies

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

In the Department of Electrical and Computer Engineering

University of Saskatchewan

Saskatoon SK, Canada

By

RASHEDUL HASAN

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other uses of materials in this thesis in whole or part should be addressed to:

Head of the Department of Electrical and Computer Engineering
57 Campus Drive
University of Saskatchewan
Saskatoon, Saskatchewan
Canada, S7N 5A9
OR
Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9
Canada

ABSTRACT

Modern artificial intelligence and machine learning opens up new era towards video surveillance system. Next generation video surveillance in *Internet of Things (IoT)* environment is an emerging research area because of high bandwidth, big-data generation, resource constraint video surveillance node, high energy consumption for real time applications. In this thesis, various opportunities and functional requirements that next generation video surveillance system should achieve with the power of video analytics, artificial intelligence and machine learning are discussed. This thesis also proposes a new video surveillance system architecture introducing *fog computing* towards IoT based system and contributes the facilities and benefits of proposed system which can meet the forthcoming requirements of surveillance. Different challenges and issues faced for video surveillance in IoT environment and evaluate fog-cloud integrated architecture to penetrate and eliminate those issues.

The focus of this thesis is to evaluate the IoT based video surveillance system. To this end, two case studies were performed to penetrate values towards energy and bandwidth efficient video surveillance system. In one case study, an IoT-based power efficient color frame transmission and generation algorithm for video surveillance application is presented. The conventional way is to transmit all R, G and B components of all frames. Using proposed technique, instead of sending all components, first one color frame is sent followed by a series of gray-scale frames. After a certain number of gray-scale frames, another color frame is sent followed by the same number of gray-scale frames. This process is repeated for video surveillance system. In the decoder, color information is formulated from the color frame and then used to colorize the gray-scale frames. In another case study, a bandwidth efficient and low complexity frame reproduction technique that is also applicable in IoT based video surveillance application is presented. Using the second technique, only the pixel intensity that differs heavily comparing to previous frame's corresponding pixel is sent. If the pixel intensity is similar or near similar comparing to the previous frame, the information is not transferred. With this objective, the bit stream is created for every frame with a predefined protocol. In cloud side, the frame information can be reproduced by implementing the reverse protocol from the bit stream.

Experimental results of the two case studies show that the IoT-based proposed approach gives better results than traditional techniques in terms of both energy efficiency and quality of the

video, and therefore, can enable sensor nodes in IoT to perform more operations with energy constraints.

ACKNOWLEDGEMENTS

I would like to express my deep and sincere gratitude to my supervisor, Professor Khan A. Wahid for supervising my work. I owe to him for his constant supervision, encouragement, personal guidance during the progress of my thesis. Starting with a little background in Internet of Things, I was able to acquire content knowledge and contribute to the advancement of the state-of-the-art research with his valuable guidance and continual encouragement. I am privileged to have the opportunity to work under his supervision, which immensely enriched my graduate experience.

I would also like to thanks all the members of Multimedia Processing and Prototyping laboratory at the University of Saskatchewan. Finally, I would like to express my deepest gratitude and love to my family and friends for their unconditional love, care, and support at each step of my life.

TABLE OF CONTENTS

PERMISSION TO USE i
ABSTRACT iii
ACKNOWLEDGEMENTS v
TABLE OF CONTENTS vi
LIST OF TABLES x
LIST OF FIGURES xi
ABBREVIATIONS AND SYMBOLS xiii
Chapter 1 - Introduction 1
1.1 Overview of Video Surveillance and Internet of Things (IoT) 1
1.2 IoT Applications 2
1.3 Objective of the Thesis 4
1.4 Organization of the Thesis 5
Chapter 2 - Literature Review 6
2.1 Video Surveillance 6
2.1.1 Evolution of Video Surveillance Systems 6
2.1.2 Intelligent Video Surveillance (IVS) 8
2.1.3 Sectors of Application 9
2.2 The Internet of Things (IoT) 10
2.3 CloudIoT 10
2.4 Fog Computing 12
2.4.1 Fog Computing in Literature 14
2.5 Internet of Video Things (IoVT) 15
2.5.1 IoVT Video Surveillance Experiments 15
2.6 Color Frame Reproduction 17
2.7 Summary 17
Chapter 3 - Next Generation Video Surveillance Requirements 18
3.1 Video Surveillance Requirements – Part 1 18
3.1.1 Video Surveillance Basics 19
3.1.2 Anomaly Detection 19
3.1.3 Surveillance Node Requirements 20

3.1.4 Camera Position Inspection 21
3.1.5 Bandwidth Allocation 21
3.1.6 Local Processing and Local Storage 21
3.1.7 Interconnectivity with Other Systems 21
3.1.8 Scalability 22
3.1.9 Vendor Independent System 22
3.2 Video Surveillance Requirements – Part 2 23
3.2.1 Reduce Manual Monitoring 23
3.2.2 Real Time Analysis 23
3.2.3 Motion Pattern Recognition 24
3.2.4 Behavior and Event Analysis 24
3.2.5 Cooperative and View Selection 24
3.2.6 Integration and Statistics 24
3.2.7 Learning and Classification 24
3.2.8 3-D Sensing 25
3.2.9 24/7 Data Goldmine 25
3.3 Example with Bank Scenario 25
3.4 Summary 26
Chapter 4 - IoT Based Video Surveillance System 27
4.1 New Era of Fog Computing Towards Video Surveillance 27
4.2 Challenges in IoT based video surveillance 29
4.2.1 Latency Requirements 29
4.2.2 Network Bandwidth Constraint 30
4.2.3 Resource Constraint Video Surveillance Node 30
4.2.4 Time Critical Video Surveillance System 31
4.2.5 Uninterrupted services Due to Intermittent Connectivity to The Cloud 31
4.2.6 Security Challenges 32
4.3 Advantages of Fog Computing in Video Surveillance 34
4.3.1 Low latency 35
4.3.2 Network Bandwidth 35
4.3.3 Geographical Distribution 35
4.3.4 Large Number of Nodes 36

4.3.5 Real-Time Interactions 37
4.3.6 Heterogeneity 37
4.3.7 Interoperability and Federation 37
4.3.8 Security 37
4.3.9 Scenario-Based Example 38
4.4 Summary 40
Chapter 5 - Case Study 1: Frame Sampling Technique in Bandwidth Constraint IoT Environment 41
5.1 Overview 41
5.2 System Design 42
5.2.1 Encoder Design 43
5.2.2 Decoder Design 44
5.3 Experiment in Different Color Space 47
5.4 Result and Analysis 48
5.4.1 Savings in Transmission Energy 48
5.4.2 Bandwidth Savings 49
5.4.3 Video Quality for different Color Space 50
5.4.4 Rate Quality Characteristics 51
5.5 Findings 53
5.6 Summary 53
Chapter 6 - Case Study 2: Frame Reproduction Technique by Variable Length Pixel Encoding 55
6.1 Overview 55
6.2 System Design 56
6.2.1 Creation of Bit Stream 57
6.2.2 Extract Pixel Information in Cloud Side 61
6.3 Experiment for Bayer (RGGB) Video Image 63
6.4 Result and Analysis 64
6.4.1 Video quality 64
6.4.2 Compression 66
6.4.3 Rate quality characteristics 68
6.4.4 Bandwidth savings 69
6.5 Advantages Towards IoT and IoVT 69
6.6 Summary 70

Chapter 7 - Conclusion and Future Work 71
7.1 Overview 71
7.1 Future Works 72
REFERENCES 74

LIST OF TABLES

Table 6.1: Bit save for different number of channel in bitstream 66
Table 6.2: Number of channels in bitstream and needed bits per pixel 67
Table 6.3: Bandwidth savings and required bits/pixel for Bayer images 67

LIST OF FIGURES

Figure 1.1: Applications of Internet of Things 3
Figure 2.1: Evolution of video surveillance 7
Figure 2.2: Traditional video surveillance setup 8
Figure 2.3: Application scenarios by CloudIoT paradigm and challenges 11
Figure 2.4: Cloud-Fog-Sensor three-tier architecture 13
Figure 3.1: Video Surveillance System in integrating with smart building components 22
Figure 4.1: Fog based distributed video surveillance system 28
Figure 4.2: Fog-cloud collaborative services for video surveillance system 28
Figure 4.3: Multidisciplinary Units of Fog based surveillance system 29
Figure 4.4: Layered architecture of the proposed IoT based video surveillance framework for power efficient, low latency and high throughput analysis of the surveillance big data 34
Figure 4.5: Multi Fog based distributed video surveillance system 36
Figure 4.6: Sequence Diagram: Responsibility of video surveillance system in Fog-Cloud collaborative design towards a robbery detection scenario of a bank 39
Figure 5.1: Proposed Vs Conventional Technique; Encoder sending a consecutive number of single component frames (Y) between three color frames in YCbCr color space. 42
Figure 5.2: Encoder sending a consecutive number of gray frames 43
Figure 5.3: Flowchart for Decoder Technique 46
Figure 5.4: Result of the proposed pre-processing operation. (a) Original frame image Y component. (b) MOTION BLOCK of the corresponding frame 46
Figure 5.5: Result of the activity level of MB. (a) Activity level of SKIP BLOCKs (b) Activity level of MOTION BLOCKs 47
Figure 5.6: Savings in transmission energy consumption for different Ep in YCbCr 49
Figure 5.7 Performance Comparison of required transmission bandwidth for different numbers of Ep in YCbCr. 50
Figure 5.8: PSNR (dB) comparison for different color space 51
Figure 5.9: MSE (0-255) comparison for different color space 51
Figure 5.10: Rate-quality characteristics of the proposed technique in YCbCr 52

Figure 5.11: Rate-quality characteristics of the proposed technique and MJPEG for different color spaces 52
Figure 6.1: Flowchart of the proposed variable length pixel encoding 55
Figure 6.2: Block diagram of the proposed variable length pixel encoding 57
Figure 6.3: Flowchart for adding a pixel information into bitstream 59
Figure 6.4: Example of bitstream creation from an image 60
Figure 6.5: Flowchart for extracting pixel information from bitstream 62
Figure 6.6: Flowchart for adding a pixel information for Bayer image pixel 63
Figure 6.7: PSNR(dB) for proposed technique with RGB images 64
Figure 6.8: MSE (0-255) for proposed technique with RGB images 65
Figure 6.9: PSNR (dB) and MSE (0-255) for Bayer images 65
Figure 6.10: Rate-quality characteristics of the proposed technique for RGB color images 68
Figure 6.11: Performance comparison of required bandwidth for different value of threshold ... 69

ABBREVIATIONS AND SYMBOLS

AI	Artificial Intelligence
DC	Data Center
DHT	Distributed Hash Table
Ep	Energy Parameter
EU	End User
FD	Frame Difference
IoT	Internet of Things
IoVT	Internet of Video Things
ISP	Image Signal Processing
IVS	Intelligent Video Surveillance
LTE	Long Term Evolution
MB	MacroBlock
MSE	Mean Square Error
NIST	National Institute of Standard and Technologies
P2P	Peer-to-Peer
PSNR	Peak Signal-to-noise Ratio
PTZ	Pan-Tilt-Zoom
QoS	Quality of Service
RA	Reference Architecture
RFID	Radio Frequency Identification
ROI	Region of Interest
SMS	Short Message Service
TN	Terminal Node
ToF	Time-of-flight
WAMI	Wide Area Motion Imagery

Chapter 1 - Introduction

1.1 Overview of Video Surveillance and Internet of Things (IoT)

Video Surveillance consists of remotely monitoring public or private places, mostly using video cameras that transmit the video/images to the server or/and monitoring unit [1]. Video surveillance guides to a range of security activities like dissuasion, observation, intelligent gathering, assessment of a possible and actual incident, forensic analysis after an incident and evidentiary [2].

The Internet of things (IoT) is the inter-networking of physical devices, vehicles, buildings and other items, embedded with electronics, software, sensors, actuators, and network connectivity that enable these objects to collect and exchange data. IoT melds together physical objects, virtual objects, living beings, user interfaces and analytics all interconnected over an internet-based infrastructure. In recent years, increasing amount of video cameras have been appearing throughout, including surveillance cameras for security, maintenance, management, healthcare and many critical purposes. Video surveillance sensors and surveillance data are interesting research topics in the field of IoT mainly for the three reasons. Firstly, the transmission bandwidth of a video surveillance node is much higher, leads obstacle for the communications systems. Secondly, the video surveillance cameras are usually been installed in any place including remote areas. So there are a lot of battery-driven surveillance video sensor available in IoT-based application. But the power consumption of those video sensors is much higher for transmitting a huge amount of data which leads higher deployment and maintenance cost. Thirdly, the huge information gathered from distributed video surveillance nodes results in ultra-big data [3]. Therefore, video data filtering, processing and data mining is also a crucial task in IoT.

Fog Computing is a paradigm that extends cloud computing and services to the edge of the network [4-7]. The vision of fog computing is to enable applications on billions of connected devices in IoT environment. Similar to cloud, fog provides data, compute, storage and application services to end users. However, fog can be distinguished from the cloud by its proximity to the end users, the dense geographical distribution and its support for mobility [8]. In traditional video surveillance system, each surveillance camera/node directly sends the video data to the cloud. However, the cloud becomes a problem for bandwidth and latency sensitive video surveillance

application, which requires surveillance nodes in the vicinity to meet their delay requirements. In this thesis, fog computing is focused in IoT environment as a solution to eliminate those issues and obstacles for video surveillance systems.

In the latter part of the thesis, two case studies are presented, which can save a significant amount of transmission energy without the need for additional hardware resources. Instead of transmitting all color images, proposed IoT based approach transmit only one color image followed by a series of gray-scale images in one case study. The gray-scale images received in decoder side are then colorized by a content-aware pre-processing and motion estimation and compensation technique. In second case study, the basic idea is to transmit the information of the pixels which is dissimilar comparing to previous frame information. With this goal, a bit stream is created for every frames information and transmit to another side of the node.

1.2 IoT Applications

Though the applications of IoT still in its early stage, however, the use of IoT rapidly evolving and growing. Depending on the application, for achieving the cost and benefit, designers have to change the design and goals [23]. Figure 1.1 shows some of the main domain of application of IoT. Below are some of the key IoT applications available in the industry.

- IoT Healthcare service [24] – IoT provides great opportunities to improve healthcare [25] by the help of IoT's ubiquitous identification, sensing and communication facilities. The devices and objects of healthcare system can be monitored and act accordingly [26]. By using the portable computing devices (laptop, mobile, tablet) and modern internet communication technology (5G, LTE, etc), IoT based healthcare services can be mobile and personalized [27]. The utilization of wearable sensors, together with appropriate applications running on individualized computing gadgets empowers individuals to track their everyday exercises (steps strolled, calories consumed, practices performed, and so on.), giving proposals for improving their way of life and keep the beginning of medical issues [28].

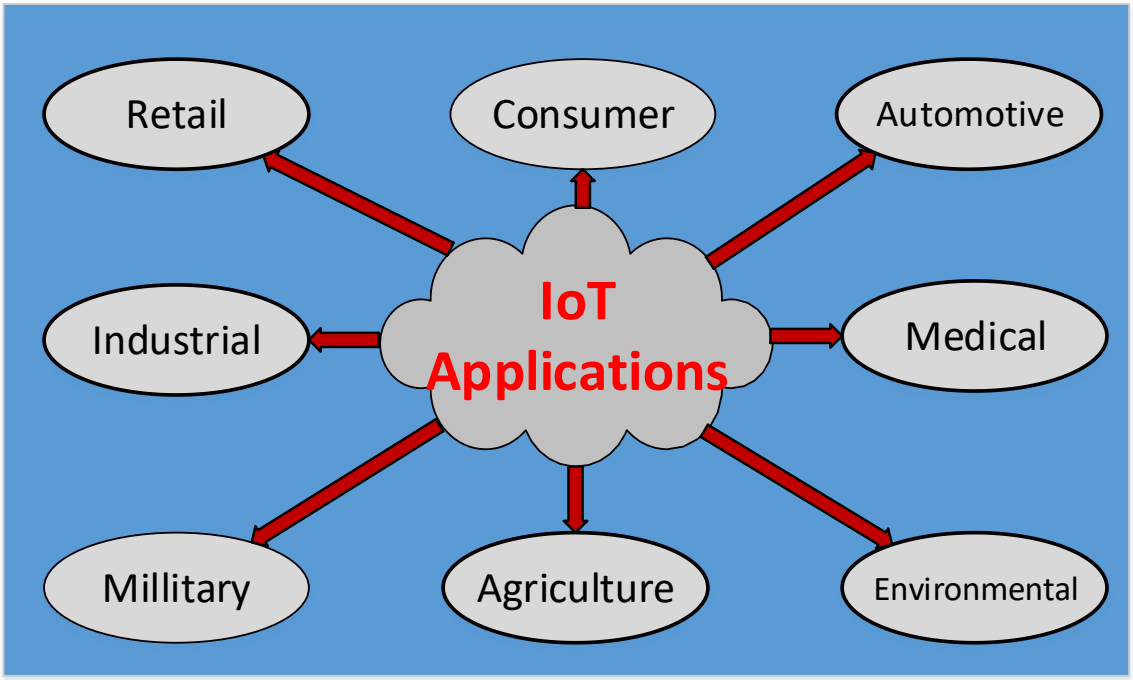


Figure 1.1: Applications of Internet of Things

- Smart Cities – As indicated by Pike Research on Smart Cities [28], the Smart City showcase is evaluated at several billion dollars by 2020, with a yearly spending coming to almost 16 billion. This market combined interconnection of key industry and benefit divisions, for example, Smart Governance, Smart Mobility, Smart Utilities, Smart Buildings, and Smart Environment. Monitoring structural health of historical building and identification of the areas that are most subject to the impact of external agents needs the support of IoT. A deeper penetration of IoT solution in Waste Management domain may results in significant savings and advantages [29]. The other areas of Smart cities which required the penetration of IoT includes measurement of air quality [30], monitoring of noise [31], traffic congestion [32], measurement and management of city energy consumption, advancement of smart parking [33], smart lighting and automation of public buildings.
- Smart business/Inventory and product management – RFID advancements are as of now utilized as a part of numerous areas for inventory management, supply and delivery chain. IoT can also play a role in after-market support, whereby users can automatically retrieve all data about the products. Different technologies of identification can help to detect thefts

and products with unique identification provides a complete and trustworthy description of the good itself. Bio-sensor technology with a combination of RFID technology may allow to identify and track the product and monitor parameters such as bacterial composition of a food product to qualify the product.

- Security and Surveillance – Security observation has turned into a need for big business structures, shopping centers, manufacturing plant floors, auto parks and numerous other open places. IoT-empowered innovations can be utilized to significantly improve the performance of current arrangements, giving less expensive and less intrusive alternatives to the widespread deployment of cameras while in the meantime saving users privacy.
- Transportation and Logistics – IoT can play increasingly demanding as well as an important role in transportation and logistics industries [34]. Physical objects with a barcode, RFID tags and sensors contribute real-time monitoring of the movement of physical objects through the entire supply chain including manufacturing, shipping and distribution [35]. Through IoT technologies, it is possible to track each vehicle's existing location and monitor its movement. , an intelligent informatics framework (iDrive system) developed by BMW, utilizing various sensors and tags to track the vehicle location and the road condition and provides driving directions [36].

Besides the above applications of IoT, the scope of IoT is extremely wide. ICT businesses, institutionalization bodies and policymakers have attempted a progression of activities to guide the IoT advancement process with the goal of augmenting its financial esteem while limiting the dangers related to security and privacy of information.

1.3 Objective of the Thesis

The main target of this thesis is to contribute towards future IoT based video surveillance system. To achieve this goal, the specific objectives of this thesis are as follows:

- To figure out the intelligent requirements of next generation video surveillance system.
- To figure out the challenges for IoT based video surveillance system.
- To demonstrate a fog based video surveillance system to eliminate the issues of IoT based video surveillance system.

- To provide an energy and bandwidth efficient frame sampling technique for IoT based video surveillance system as a case study.
- To provide an energy efficient variable length pixel encoding technique for surveillance applications as a case study.

1.4 Organization of the Thesis

A brief overview of the organization of the rest of the thesis is presented in this section. In chapter 2, traditional video surveillance system and its evolution are discussed. The Internet of Things and fog computing in literature are rigorously reviewed here. In chapter 3, intelligent requirements that should be incorporated in next generation video surveillance system are presented. The challenges of IoT based video surveillance system and the proposed fog-cloud integrated architecture for video surveillance are discussed in chapter 4. This chapter also includes the opportunities of fog computing and how fog computing can be a solution to overcome the challenges in IoT based video surveillance system. In chapter 5, a color frame transmission and reproduction technique that is applicable for energy and bandwidth efficient IoT based video surveillance system is presented. In chapter 6, a variable length pixel encoding technique for video frame transmission is presented which is also applicable for IoT based video surveillance system. Finally, conclusion and direction for future research are given in chapter 7.

Chapter 2 - Literature Review

2.1 Video Surveillance

The objective of video surveillance consists of remotely monitoring objects such as public or private places by video cameras that transmit the video/images to the server/cloud or/and monitoring unit [1]. It has been widely adopted in various cyber-physical systems including traffic analysis, healthcare, public safety, wildlife tracking, smart building, industry automation and environment/weather monitoring. For example, the traffic signal system developed by the transportation department in the city of Irving, Texas [9], the traffic monitoring system at the University of Minnesota [10] and the system at the University of North Texas [11] are some examples of traffic video surveillance which were implemented more than a decade ago. Remote weather monitoring system (FireWxNet, 2006) in the Bitterroot National Forest in Idaho to monitor the lightning stricken forest fire [12], the smart camera network system (SCNS, 2011) used for security monitoring in a railway station [13], and the indoor surveillance system in a multi-floor department building at the University of Massachusetts-Lowell [14] are the video surveillance application examples of different areas.

Video surveillance camera or surveillance node mainly consists of video sensor, encoder, processing unit and transmitter. Both wireless and wired video surveillance nodes are used in surveillance system. Wired surveillance node has advantages of having no interference from nearby nodes and stable data transmission. But it is difficult to install wired video surveillance node and require drilling holes. The wired node is fixed in place after installation. Wireless surveillance node has advantages of easy and quick installation, flexibility and portability. Video surveillance nodes are both managed and un-managed. Un-managed nodes has only one way communication from node to router/server. The configuration of un-managed nodes usually set before the installation. It does not support remote configuration management system. On the other hand, managed video surveillance node has two way communication system. The configuration of managed surveillance nodes can be changed remotely in run-time.

2.1.1 Evolution of Video Surveillance Systems

Transition in video surveillance occurred in several steps. It began with the arrival of the digital recorder, then continued with the modernization of the IP infrastructure, with the video

being transmitted over an intranet or internet network from the camera. Many hybrid systems took shape during this shift, integrating analog and digital components. Figure 2.1 shows the evolution towards intelligent video surveillance. This evolution can be divided into three generations. In the first generation, analog cameras are connected by coaxial cables to surveillance screen and for archiving purposes, to a videotape recorder that records the video on a cassette. Second generation video surveillance replaced the videotape recorder with a digital recorder. In the third generation, all of the video surveillance components are digital and all transmissions are done by IP protocol. Those networks, therefore, have network cameras or IP cameras which have their own built-in encoder. Those cameras are connected to servers (PCs) equipped with video management software. Videos are stored on a server or on proprietary network video recorders [1]. A traditional video surveillance system setup is shown in Figure 2.2. Infrastructure and principles of video surveillance system mostly integrated with acquisition, transmission, compression, processing, archiving and display [15, 16].

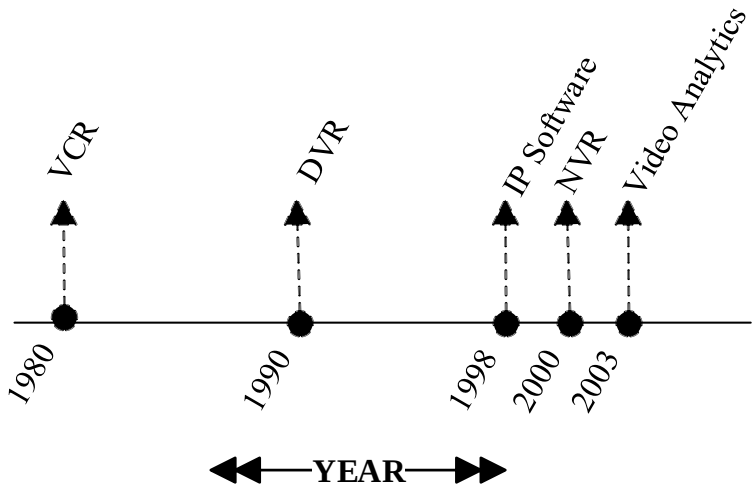


Figure 2.1: Evolution of video surveillance

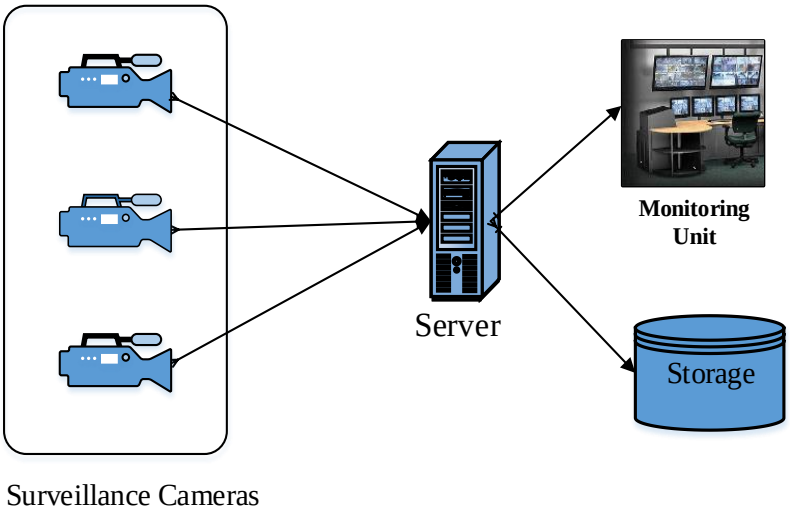


Figure 2.2: Traditional video surveillance setup

2.1.2 Intelligent Video Surveillance (IVS)

Intelligent video surveillance (IVS), is a technology that uses software to automatically identify specific objects, behaviors or attitudes in video [1]. It transforms the video into data and video surveillance system can act accordingly. It may involve sending a warning to surveillance personnel so that a decision may be made on the proper intervention. Intelligent video surveillance systems use mathematical algorithms, machine vision to detect moving objects in an image and filter necessary objects.

IP camera and server/cloud based video surveillance along with the massive improvement of video analytics exposed a new era of intelligent video surveillance during the last decade. According to D. Elliote in [17], intelligent video surveillance is defined as “any video surveillance solution that utilizes technology to automatically, i.e., without human intervention, process, manipulate and/or perform actions to or because of either the live or stored video images”. The increasing demand in this area challenges both academic researchers and industrial practitioners to provide analytics theory and system solutions timely to meet the overwhelming global need. In 2012, the video analytics market had been about 60% annual compound growth [18].There are over 6000 research papers published since 1971 in video systems design, tracking, modeling, behavior understanding, abnormal detection, real-time performance and practical implementation of intelligent video systems and analytics [18].

2.1.3 Sectors of Application

Previously video surveillance was only used by public services (police, transportation, administration). It was then used by companies looking to protect assets, such as power plants, agro-food and different complexes. Casinos also appear as pioneers in the distribution of large video surveillance system. Today, surveillance cameras can be seen in different public and private places, such as apartment building, parking lots, bus/train station, roads, banks and shops etc. In recent era, video surveillance is used in mobile units, such as patrol cars, ambulances, buses etc.

1. Government and Public Security - Video surveillance system is used to monitor sensitive infrastructures, borders, government building, laboratories, military bases and prisons on the national level. It is important in several cities around the world to monitor crimes and as an emergency intervention tool. It is also used to ensure security in large gathering (shows, demonstration and sporting event). It is used for parking management for monitoring parking permits, application of rules, theft detection, vandalism and access control system.

2. Education - Video cameras are widely used in academic institutions. It is used to monitor the safety of students and asset protection from theft and vandalism. In the university environment, video surveillance is used in particular to:

- Monitor access to institution's perimeter, which may be extensive, such as a university campus
- Monitor important data, tools, devices and equipment
- Detect misbehavior, ill motive, theft and vandalism.
- Detect license plate in the parking lot.
- Access control in the secured area.

3. Retail Trade - Video surveillance is used for internal (store, warehouse) and external (parking lot) security of retail trade. In chain shops, sophisticated video surveillance system is set up for monitoring centrally of different locations. For this entire sector, video surveillance is directed primarily at:

- Monitoring transactions and resisters (employee theft and fraud).
- Protecting infrastructure and material goods.
- Monitoring deliverables and inventory.

- Controlling access to secured and locked areas.
- Checking and acting in the emergency situation.

4. Transportation - Smooth operation of airport, railroads/bus station, ports and mass transit system heavily depends on video surveillance system. It is used to detect an intrusion in a controlled area, detecting people entering or exit, luggage, recognize faces, counting people, recognize license plates.

5. Bank Setting - Video surveillance is widely used for bank environment for their security. Automated bank machines or ATMs are prime targets for criminal acts. Surveillance cameras help the bank to detect fraud, such as, for example, the installation of a device to read the magnetic information on bank cards. In a bank setting, intelligent video surveillance can increase monitoring effectiveness and increase the efficiency. It provides for monitoring of all branches to detect suspicious behavior.

2.2 The Internet of Things (IoT)

The Internet of things (IoT) is the inter-networking of physical devices, vehicles, buildings and other items, embedded with electronics, software, sensors, actuators, and network connectivity that enable these objects to collect and exchange data [19]. IoT melds together physical objects, virtual objects, living beings, user interfaces and analytics all interconnected over an internet-based infrastructure. As identified by Atzori et al. [20], Internet of Things can be recognized in three paradigms—internet-oriented (middleware), things or object oriented (sensors) and semantic-oriented (knowledge). As a promising technology, IoT offers promising solutions to transform the operation and role of many existing industrial systems and manufacturing systems. The term IoT was first proposed to refer to uniquely identifiable connected objects with radio-frequency identification (RFID) technology [21]. However, in the past decade, the definition has been more inclusive covering varieties of applications like healthcare, utilities, transport, etc. In 2011, the number of interconnected devices overtaken the actual number of people in the world and it is expected to reach 24 billion interconnected devices by 2020 [22].

2.3 CloudIoT

Cloud computing [37] is the most recent paradigm in this era which promises reliable services delivered through next generation database and data centers that are based on virtualized

storage technologies. Cloud computing is internet-based computing, where resources, software and information are shared among computers and devices. The cloud computing concept is highly matured over the last few years. The concept reveals that anything hosted on the internet can be available for use when needed for more sophisticated services [38]. Few cloud components are on-demand service provision, ubiquitous access, resource pooling and elasticity. Cloud computing has virtually enormous capabilities with respect to storage and processing power. The fundamental aspects of Cloud computing was reported by National Institute of Standard and Technologies (NIST) [37] ‘‘Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction’’.

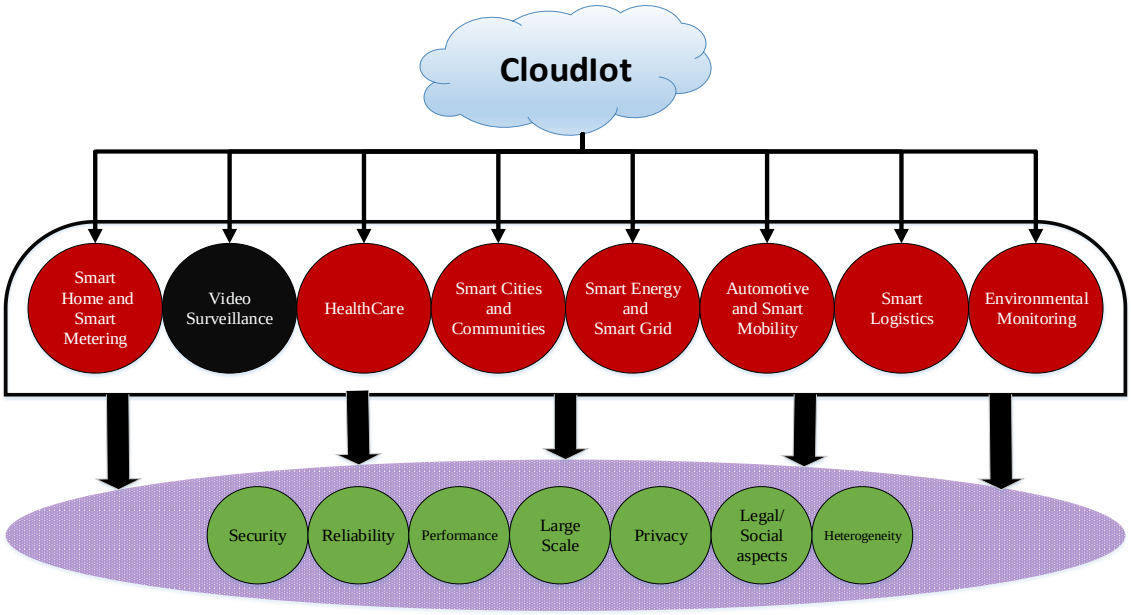


Figure 2.3: Application scenarios by CloudIoT paradigm and challenges

Cloud computing provides extreme opportunity to solve most of the IoT issues. A novel IT paradigm was proposed which is called CloudIoT, merging these two complementary technologies [39, 40]. IoT gets help from the virtually unlimited capacities and resources of Cloud to satisfy its technological constraints (e.g., storage, processing, and communication). Cloud offers an efficient solution for IoT, hence implementing applications and services that exploit the things or the data produced [41]. On the other hand, Cloud can benefit from IoT by dealing with real-world things

in shared and dynamic manner, and for introducing new services in a large number of real-life scenarios Figure 2.3 indicates the application of CloudIoT and related challenges.

CloudIoT helps intelligent video surveillance leads video contents originating from video sensors to easily and efficiently store, manage, and process. It also helps automatically extract knowledge from scenes. In [42], F. Gao proposed solutions that are able to deliver video streams to multiple devices through the Internet, by distributing the processing tasks over the cloud on-demand, in a load-balanced and fault-tolerant fashion. Complex video analytics needs Cloud-based solutions [43] to properly satisfy the requirements of video surveillance system (e.g., stored media is centrally secured, fault-tolerant, on-demand, scalable, and accessible at high-speed) and video processing (e.g., computer vision and machine learning module execution).

2.4 Fog Computing

Fog Computing is a paradigm that extends cloud computing and services to the edge of the network [4-7]. The vision of fog computing is to enable applications on billions of connected devices in IoT environment. Cloud computing suffers from substantial end-to-end delay, traffic congestion, processing of huge amount of data and communication cost due to the significant physical distance between cloud service provider's Data Centers (DCs) [44] and End User (EU). Fog computing introduces in order to penetrate those issues as an alternative to traditional cloud computing to support latency sensitive, geographically distributed and Quality of Service (QoS)-aware IoT applications. Fog computing was first commenced by Cisco to expand the cloud computing to the edge of the network [4, 45]. Fog computing is a highly virtualized platform [46] that can serve computing, storage and networking between EU and DC of the traditional cloud computing. However, fog can be distinguished from the cloud by its proximity to the end users, the dense geographical distribution and its support for mobility [8].

The OpenFog Consortium [47], is a consortium of large companies and academic institutions, which aims to standardize and promote fog computing in IoT applications. This consortium was established by some giant companies like ARM, Cisco, Dell, Intel, Microsoft Corp. and Princeton University Edge laboratory on November 19, 2015. Very recently, OpenFog announced its Reference Architecture (RA) for Fog Computing on 13 February 2017 [48]. OpenFog Consortium workgroups are working towards making an open standard for Fog processing to empower interoperability and adaptability. Moreover, Fog computing is also

supported by some organizations, for example, Cloudlet [49] and Intelligent Edge by Intel [50]. Chiang and Zhang [51] discussed many technology enablers for fog computing in various fields. The flow of current research reflects the enormous potential of fog computing towards sustainable development in world-wide IoT market.

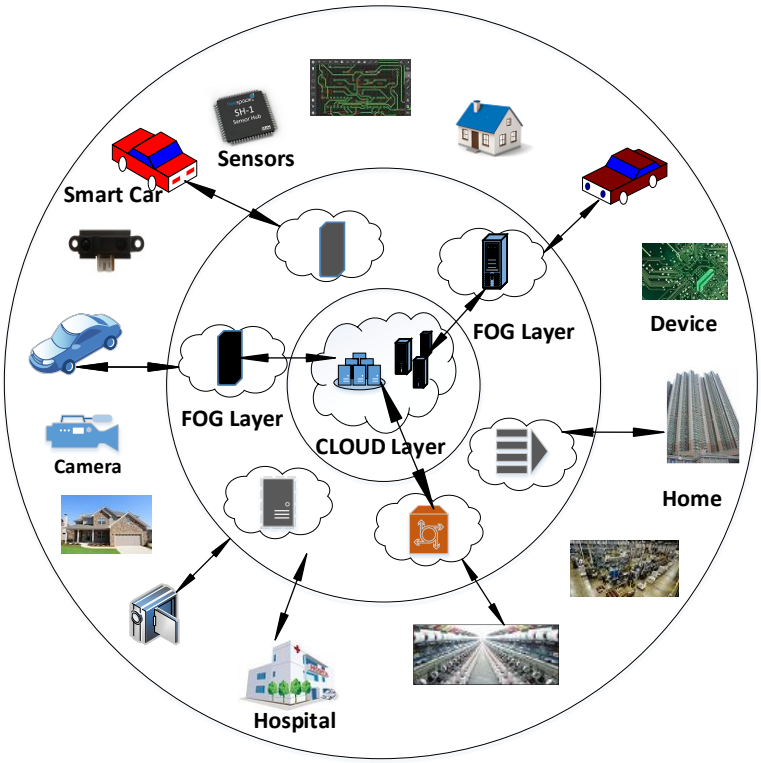


Figure 2.4: Cloud-Fog-Sensor three-tier architecture

One of the basic and widely used architectures of Fog computing is three-tier architecture [52, 53]. Figure 2.4 illustrates the three-tier architecture. The bottom tier consists of IoT-enabled devices and sensor nodes. End user's smart hand-held devices (smartphones, tablets, wearable devices, smart watches) are also in this tier. These end devices are often named as Terminal Nodes (TNs). The middle tier is also named as Fog computing layer. The fog nodes in this layer are collections of networking devices such as switches, gateway, routers etc. These fog nodes can co-operatively share storage and computing facilities. The upper tier consists of traditional cloud server and cloud data centers. This tier usually has sufficient storage and computing resources.

2.4.1 Fog Computing in Literature

Recently, many articles have investigated the state of the art of fog paradigm and examined the promising IoT applications in varieties of areas, for example, smart city, smart home, smart grid and smart healthcare.

- In [7], Yi et al. discussed the definition and representation of fog application scenarios, which includes content delivery and caching, real-time video analytics and mobile big data analytics in the context of fog networking, computation offloading and resource management.
- The theoretical modeling and compared performance of Fog computing with cloud computing in terms of service latency and energy consumption were discussed in [54].
- In [55], Varghese et al. highlighted the feasibility and compared with the cloud-only model. This research revealed that the fog computing can reduce the average response time by 20% for a user and the data traffic by 90% between the network edge and the cloud.
- The suitability of fog computing was discussed in [52] by comparing power consumption, service latency, CO2 emission and cost, and evaluating its performance for an environment with a large number of end-devices. The authors also showed that fog computing can decrease by 50% on response delay compared with the traditional cloud computing.
- Dantu et al. in [56], feasibility for the smartphones as fog nodes was analyzed and the reliability and adaptability of deploying fog computing on Android phones were discussed.
- The use cases of fog computing in surveillance, smart power grid and drones are discussed in [57], and the distinguished features of fog, edge and cloud computing on resource characteristics, physical access and mobility support based on the feasibility of deployment was compared.
- Perera et al. in [58] reviewed the use case scenarios in smart cities and evaluate the common features of fog computing, including dynamic discovery, management and configuration of Internet objects, network-level protocols and application-level protocols.
- 5G ecosystem to support smart cities and industrial automation was discussed in [59].
- Some articles [60 - 62] examined the smart gateways at the edge of the network to offer many high-level services, such as storage, real-time processing and data analysis, based on fog computing and presented novel system in smart healthcare.

- Roman et al. in [63], an extensive analysis of the security threats, challenges and mechanisms was presented to explore potential synergies and venues of collaboration on edge paradigms, including fog computing, mobile edge computing and mobile cloud computing.

2.5 Internet of Video Things (IoVT)

Among different types of Internet of Things system, with the rapid development of communication, computation ability and computer vision technologies, Internet-of-Video Things (IoVT), where video cameras are deployed as the major sensors of the IoT. IoVT has a high potential for wide IoT applications since rich context information can be emitted from video data. This distinction is needed and necessary because IoVT has very different requirements than traditional IoT devices.

In recent years, increasing amount of video cameras have been appearing throughout, including surveillance cameras for security, maintenance, management, healthcare and many critical purposes. Video surveillance sensors and surveillance data are interesting research topics in the field of IoVT mainly for the three reasons [64]:

1. The transmission bandwidth of a video surveillance node is much higher, leads obstacle for the communications systems.

2. The video surveillance cameras are usually been installed in any place including remote areas. So there are a lot of battery-driven surveillance video sensor available in IoT-based application. But the power consumption of those video sensors is much higher for transmitting a huge amount of data which leads higher deployment and maintenance cost.

3. The huge information gathered from distributed video surveillance nodes results in ultra-big data [3]. Therefore, video data filtering, processing and data mining is also a crucial task in IoVT.

2.5.1 IoVT Video Surveillance Experiments

One of the fundamental goals of this thesis is to penetrate value on energy and bandwidth efficient IoVT-based video surveillance system. Some prior researches have been conducted on the same objective.

For video surveillance system in IoT paradigm, slot allocation of multi-camera video streaming that is validated under IEEE 802.15.4e wireless personal area networks was used in [65]. Pereira et al. in [66], a list of possible recommendations for low computation and memory capacity of a video surveillance node, connected via low-quality wireless channels was discussed. He et al. in [67], a power-scalable video encoding method was proposed to minimize the energy consumption in portable video communication devices.

Malic et al. in [68], an IoVT video streaming experiments was conducted where open source hardware, NoSQL database, Cloud and IoT platform were used. The captured images and videos were transferred to a remote server which acted as a cloud. They used MongoDB for document storage and Arduino Yun with OpenWRT-Yun as operating system. 640x480 pixels resolution was used and the experiment succeeded in streaming video wirelessly to a web server.

Feasibility of IoVT video streaming is measured with respect to overhead and delay in [69]. In this experiment, Peer-to-Peer (P2P) network and "sensible Things" platform for IoT network layer were used. Distributed Hash Tables (DHT) was used while experiment conducted with P2P. H.264 encoding and decoding units on the raspberry pi platform were used due to availability. To minimize the delay of video transmission, this approach allowed to send minimal chunks of encoded videos as small as P2P packets. The video encoding processes consumed more than 90% of the total delay even with these enhancements.

To decrease the transfer time and off-loading to a cloud based service, multiple solutions were investigated. A fog computing platform with urban traffic surveillance was used as a case study [70]. To meet the requirements of real-time video processing, a dynamic video stream processing scheme is proposed. Fog computing was used to tackle the transfer-delay problem associated with cloud computing. Fog also used to provide more processing power to the IoT devices. Wide Area Motion Imagery (WAMI) which is characterized by its high data rates and wide area coverage was used. The experiment used the technique to transfer only a sub-area of the image to a processing unit, not the whole image. The sub-area was calculated by the fog unit which provides more computation capable unit. This fog-based architecture allows for video streams to be processed on the local fog to minimize delay and only send metadata to control unit.

To accelerate the deployment of IoT services, another approach is to use the current establishment. In [71], existing image processing pipelines and modules from mobile devices were

used. The IPP extracts and pre-processes image frames prior to compression. De-mosaicking and multi-frame operations were performed as preprocessing. Image signal processing unit (ISP) were used for those operations.

2.6 Color Frame Reproduction

In conventional video surveillance system, color images are captured by the camera sensor and then transmitted out through a wireless channel to a server where the images are stored. Transmitting all three color components (i.e. red, green and blue) of a color image consumes significant processing time and battery power for a battery-driven surveillance node. There are many color generation algorithms found in the literature that mainly provides color tone from a color image to a gray-scale image [72].

Welsh et al. in [73], the color information is assigned to each pixel in the gray image by matching each gray-scale image pixel to a pixel in the target swatch using the Euclidean distance matric. In [74], Horiuchi et al. have used a set of seed points and their respective color vectors in the RGB format with a YUV-based classification. In [75], a quadratic objective function-based optimization method is used to interpolate the U and V components of the YUV color space over the entire image using a set of color scribble lines. In [76], pseudo colors are employed to colorize the gray-scale image using different 64 x 3 color matrices. In [77], an adaptive dictionary-based color generation scheme that is targeted for capsule endoscopy application. The original color tone can be reproduced from a theme image without introducing additional color artifact. In [78], an energy-efficient wireless video sensor node with content aware pre-processing and an energy and content aware feedback control scheme were used.

2.7 Summary

In summary, in this chapter, Video Surveillance System was discussed along with literature review. Internet of Things and fog computing, a new dimension of IoT has also been discussed. Internet of Video Things in IoT paradigm and some IoVT experiments in literature have been reviewed. At the end, some color frame reproduction techniques in literature were discussed. In next chapters, next-generation video surveillance requirements and the challenges in IoT based video surveillance system will be presented.

Chapter 3 - Next Generation Video Surveillance Requirements

Market trends and recent technologies have demanded the significant need for advantageous solutions to video surveillance system and analytics. At this moment, current video surveillance might not as good as people expect, however, the application of technology does spread rapidly. This chapter provides an extensive account of modern and future intelligent video surveillance requirements. It highlights the tasks, method and related video analytics that can meet next-generation surveillance requirements.

This chapter presents design requirements for IoT-based next generation video surveillance system. To illustrate the challenges and requirements of next generation video surveillance system, a scenario of a bank is considered throughout this and the subsequent chapters. In this scenario, the bank has multiple branches which are located in geographically distinct areas, maybe in different cities. Hundreds of video surveillance cameras are installed in different places of each branch of the bank. The cameras are placed in a way that every sensitive place inside the bank can be monitored. There are also some cameras placed outside the bank to monitor the entrance, doors, and windows of the bank. The requirements and the challenges of next generation video surveillance are discussed with illustrating the example of this bank scenario. For better understanding, some other video surveillance application scenarios have also been introduced which suit for detail discussion. This chapter is divided into two parts. The first part includes basic requirements of video surveillance system which mostly depends on video analytics and computer vision technology. The second part contains the upgraded requirements which mainly rely on recent advancements in machine learning, artificial intelligence and data mining.

3.1 Video Surveillance Requirements – Part 1

The future of video surveillance should meet elemental requirements as well as some higher level requirements. The traditional video surveillance system can only capture and store frames and sometimes can detect some anomalies. In this Section, several basic and higher level requirements are discussed which next generation video surveillance system should perform.

3.1.1 Video Surveillance Basics

A video surveillance system should perform the basic operation of surveillance. It should take videos from the video sensor node. Those video frames should be stored in a secured storage for a predefined time period [79].

3.1.2 Anomaly Detection

Video surveillance system should detect the anomalies depending on the application and specific system requirements [2]. For this, the intelligent video surveillance system should perform the following activities:

- Object Detection and Classification - The system should detect appearances, disappearances, and movements of objects in scenes viewed by individual nodes. Object detection can be done by various techniques such as frame differencing, Optical flow and Background subtraction [80]. In this aspect, Huang et al. in [81], introduced video arrays for real time tracking of a person, head, and face in a controlled room.

  After the object is detected, the system should classify it according to shape, color, and other properties. The approaches to classifying objects are mainly based on shape, motion, color and texture [80].
- Object Tracking - Tracking is defined as the path of an object in the image plane when it is moving. The approaches to tracking the objects are point tracking, kernel tracking and silhouette [80].
- Intra camera tracking - It can track the movement of targets over time, within fields of view of individual sensor nodes/cameras.
- Multi-camera tracking - Tracking target movements across time, from multiple cameras viewing the target simultaneously. Integration of information from multiple cameras is essential in intelligent surveillance system [82]. In [83] and [84], Wang and Comaniciu et al. discussed the development of relevant technologies for multi-camera tracking from the perspectives of computer vision and pattern recognition.
- Target-centric anomaly detection - The system should detect targets exhibit suspicious behavior, such as checking locks on all the doors along a corridor.

Let us consider the bank scenario that was described earlier. The system should detect and classify the customers, employees and the security people. For the safe and secure transaction, it

can identify a customer and can share the information with the bank transaction system. Whenever the system can detect a suspicious customer, it should track the customer. Even if the suspicious customer is out of the range of a video surveillance node, he should be tracked by the other surveillance node. By a combined effort of the surveillance nodes of different places, the system should better understand the motive of the suspicion.

3.1.3 Surveillance Node Requirements

Video cameras are the core of video surveillance system. It is responsible for capturing video frames. Along with the basic operational requirements, the camera manufacturers should meet some requirements in order to a successful deployment of video surveillance system.

- Low Powered - Surveillance Node should be low powered in terms of computation and transmission. It should be powered by an internal battery when external power is interrupted for short period of time.
- Moving Camera - Surveillance Node should be moved to the right, left, up, down or any angle. The movement should be controlled from the system core. When an anomaly is detected, movement of the surveillance node can track the moving anomaly object.

  Many recent surveillance cameras have moving facilities but in future, it should have the option to control the camera movement from the system core. In mentioned bank scenario, whenever the system can detect any suspicious behavior of a customer, with the help of intelligence, it should move the corresponding camera for a better understanding of the suspicious customer and his behavior.
- PTZ (Pan-Tilt-Zoom) - The node should be automatically activated for pan/tilt/zoom movements. It helps the surveillance system to follow objects or individuals moving along the scene or to zoom into areas of interest [85].
- Panoramic - The camera should have the facility to capture images with 360-degree visibility with a single camera and virtual PTZ in the image. Though the current resolution of those panoramic images are often not enough for analyses high level of detail. However, it is expected that future improvement of panoramic technology shall provide high resolution to analyses details [86].

3.1.4 Camera Position Inspection

The video surveillance system should identify individual camera's geometry (pan, tile or zoom) automatically. Camera position can be determined by the background image of the video captured by the camera. If a camera is not in the correct position, the system should identify it. Think about the bank scenario; someone can intentionally change the camera position to hide anything due to ill motive. In that case, the system should detect that the camera is not in the right position and can help to improve the bank security.

3.1.5 Bandwidth Allocation

Video surveillance system can measure the allocated bandwidth for the system. According to the currently allocated bandwidth, surveillance core system can change the priority, frame rate, resolution of an individual node to maintain the bandwidth. For example, in mentioned bank scenario, whenever an anomaly is detected, the core system can increase the bandwidth of the corresponding surveillance node to grab high-resolution images of the suspicion.

3.1.6 Local Processing and Local Storage

Video surveillance system has the minimal local processing power. With the local processing, it can detect any anomalies and action can be taken against it. Video Surveillance System should have local storage. There should be some mechanism by which it can select which frame or image/video information should transfer to cloud in the bandwidth-constraint environment. Other than that, frames may be stored in the local storage for archiving purposes.

3.1.7 Interconnectivity with Other Systems

Intelligent video surveillance system should communicate with the other external systems. Figure 3.1 describes some of such external systems with which surveillance system can collaborate with a smart building. It should establish a high abstraction layer communication with Alarm System, Robbery Detection System, Terrorism Detection System, Building Security System, Lift Control System, and Electronic Equipment Control System. It should act as a web client where needed and use short message service (SMS). Suppose in bank scenario, whenever the system can detect any robbery, it may communicate with the bank's central alarm system to fire it on. It should also communicate with the robbery detection system in that case. If the system can detect any terrorist activity inside or outside of the bank, it should also automatically contact with police or terrorism detection unit. It may use automatic phone call service to 911 in real time without delay.

If any fire or smoke is detected, it should communicate with the lift control system to stop running the lift. It should also use SMS or web client to communicate with the proper authority.

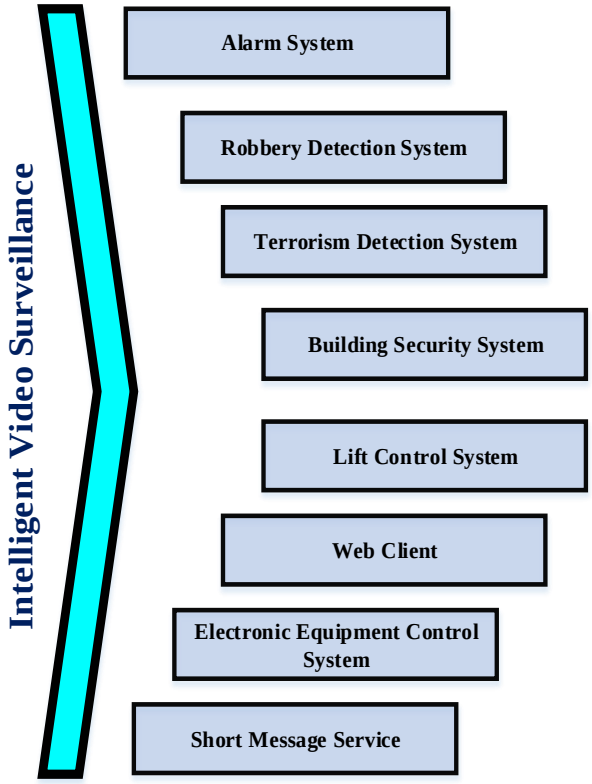


Figure 3.1: Video Surveillance System in integrating with smart building components

3.1.8 Scalability

According to the requirement, the system should add new surveillance node or can remove any surveillance node from the system. The system should also replace a surveillance node to another place. The algorithm must facilitate the replacement. For the bank scenario, it may need to rearrange the physical structure of the bank or expand the bank. In that case, the administration should change the surveillance nodes position, priority etc. Surveillance nodes should be added to system in case of expanding the bank or if new surveillance node is required.

3.1.9 Vendor Independent System

The video surveillance system should be vendor independent. So if it should be possible to replace a video node of a provider with a node of another provider with minimal inconvenience.

The core surveillance system is developed in such a way it should have no dependency with the node vendors in order to run the system.

3.2 Video Surveillance Requirements – Part 2

With the help of emerging era of video analytics and artificial intelligence (AI), the future surveillance system should perform more intelligent decision making. Video surveillance system should have the intelligence to set the priority, frame rate of individual surveillance node depending on the current situation. It should measure which nodes priority should be higher when an anomaly is detected depending on the type of detection. Some video surveillance analytics intelligence are discussed in [1].

3.2.1 Reduce Manual Monitoring

Eliminating human error is a key driver behind bringing AI to security through intelligent video analytics. Studies have shown that humans engaged in monotonous tasks have a directed attention capacity for up to 20 minutes, after which the human attention span begins to decrease. In addition, when humans are faced with multiple items at one time, attention spans will decrease even more rapidly. Therefore, video analytics should take the place of initial human judgment in an effort to increase operational efficiency. For example, a security officer of bank scenario might miss a person sneaking into a poorly lit facility, all without missing a beat and keeping a close watch on the many cameras and locations. So rather depending on human monitoring solely, AI-powered systems instead notify security teams of potential threats as they happen, helping businesses prevent break-ins or illegal activity, obviously helping accuracy.

3.2.2 Real Time Analysis

With the use of artificial intelligence and machine learning, the system should analyze the video in real time. Events should be immediately detected for triggering corresponding alerts. For this, a lot of research has been conducted in this decade, e.g., idle object detection, trajectory tracking and spatial video denoising [87] in live image sequences, a fuzzy genetic algorithm to boost the computing efficiency in a large image region [88], and a real-time system for detecting tailgating [89] are some of them.

3.2.3 Motion Pattern Recognition

Motion detection is extremely important in next generation video surveillance. Automatic interpretation of human and object motion in surveillance videos is inexorable to detect abnormal behaviors [18].

3.2.4 Behavior and Event Analysis

Surveillance cannot rely solely on the motion to identify threats, but video content analysis must be carried out for abnormal behavior detection or event recognition. Behavior and event recognition should detect and track objects in video images, applying a set of rules to detect crimes or violations such as abandoned objects, object removal, abnormal moving, crowding, loitering or stalking. Behavior and event recognition are already used in many surveillance tasks, e.g., intrusion detection [90], neurological diagnosis and management to the analysis of movement disorder [91].

3.2.5 Cooperative and View Selection

In multi-camera systems, there is a problem of choosing the right view from multiple video stream. With the blessings of artificial intelligence, efficient video understanding algorithm should be developed to automatically detect people or vehicles and seamlessly track them. Dynamic adaptation of the system is also necessary through scenarios and applications.

3.2.6 Integration and Statistics

Statistics can be applied for event detection, counting, routing, guiding, surveillance and flow control. There are already been some works for crowd flow and traffic flow on attention control and statistics for business, tourism, public safety, civilization, exhibition and market/shops [92].

3.2.7 Learning and Classification

Learning and classification are powerful ways for object detection and event recognition [89]. By the grace of machine learning, the system can be trained for detection and after learning, video surveillance system should classify threats. An adaptive learning method can be used to estimate the moving speed of a person, locating the region of interest (ROI) from an image.

Computer vision system can be inspired by the human visual system for organizing the different visual routines. For example, bio-inspired adaptive hyperspectral imaging for real-time

target tracking [93], brain-inspired neural cognitive approach for thermal image analysis [94] are some of them.

3.2.8 3-D Sensing

3-D reconstruction from 2-D images has been important for solving many problems in robotics and in many video surveillance applications. Technologies such as stereo vision [95], laser scanning and finding structure from video motion are common sensing methods to obtain 3-D [96]. Recently in [97], 3D ToF (Time-of-flight) image sensor is used to improve the vision of the things of IoT by giving a sense of depth.

3.2.9 24/7 Data Goldmine

Objects' (may be a customer) behaviors are recorded in the video and it contains valuable information for improvements in marketing effectiveness, store operation, building layout design, traffic pattern and other activities [98]. It is hard, labor-intensive and time consuming work to review hours of video from hundreds of cameras. Intelligent system should analyze a large amount of video data very quickly.

3.3 Example with Bank Scenario

The advantages that should be come out for the bank scenario by introducing intelligence in surveillance system is now figured out. By analyzing in real time, the system can identify any fraud customer or any ill behavior. Analyzing the video manually or in offline may not restrict the ill motive. It is needed to identify in real time to catch the customer/people. For this, the system may use motion pattern recognition. The system may analyze the customers' behavior and event/gesture created by the customer to identify the ill motive. For a correct understanding of the behavior of the bank customer, it may need to analyze multiple camera images to collect available views. The system may use different statistical and artificial intelligence (AI) to correctly identify the ill motive. The banks' surveillance cameras run 24 hours 7 days a week. So the system has a vast amount of previous data/video available. By intelligent data mining, learning by iteration intelligent classification algorithm, the bank video surveillance system may upgrade itself to identify a new type of behaviors. 3-D sensing also helps the system in this circumstance. Those analyses can also help to define marketing strategies. By classifying different type of bank customers by the system, the bank representative can focus the corresponding customer's interest which can generate new dimension towards profitability. Think about analyzing a new customer's

behavior and classify the customer in real time, the representative can spot the customer's class (VIP or not) on the moment before entertaining him; hence give a new dimension for customer oriented business.

3.4 Summary

This chapter has summarized the recent development and future requirements of intelligent video surveillance system. A bundle of requirements are investigated in regard to solution which involve many computer vision and machine learning tasks. In subsequent chapters will discuss how these requirements can be met in IoT based video surveillance system.

Chapter 4 - IoT Based Video Surveillance System

Internet of Things opens a new research opportunity for video surveillance system. This chapter exposes the structure of video surveillance system in Internet of Things environment. Next generation video surveillance requirements that were considered in previous chapter can be met in IoT based system. The challenges of IoT based video surveillance system are thoroughly discussed. In this chapter, fog computing is explained as an integrated part of video surveillance, and concurrently show the advantages of Fog computing for video surveillance system.

4.1 New Era of Fog Computing Towards Video Surveillance

Fog computing is an extremely virtualized platform that serves computation, storage and networking between end devices and traditional cloud servers [7]. According to [51], fog contradicts with cloud in three dimensions. Firstly, in the traditional cloud-based system, data storages are only provided in cloud servers whereas fog consists of a substantial amount of data storage at or near the end user. Secondly, cloud performs all the functions of computing and controlling in remote data centers and cellular core networks. On the other hand, fog provides applications for end users, some controlling and operation facilities, decision making, managing and also supporting cloud. Thirdly, all the networking facilities such as routing and forwarding network traffic are set by backbone network whereas fog provides a substantial amount of communication and networking at or near the end user.

In this section, use fog architecture for IoT-based video surveillance system is proposed. Figure 4.1 shows fog and IoT-based video surveillance system from the upper view. Here, surveillance camera nodes are connected to the local fog, which is then connected to the cloud.

Fog and cloud are interdependent to each other so that cloud services may be used to manage the fog. Comparing to traditional cloud-based video surveillance system, fog can act as the proxy of the cloud to deliver different analytic analysis services for intelligent video surveillance. Figure 4.2 describes conceptual fog-cloud collaborative service architecture for video surveillance system. In this architecture, Fog collects data (video frames) from surveillance nodes. All the intelligent video surveillance analytic services that have already been discussed above shall be served by fog and cloud collaboratively. Fog will be composed of a number of units for video

surveillance system. Figure 4.3 describes the units that may be integrated with fog. The subsequent sections of this chapter elaborate the challenges in IoT-based video surveillance and how fog can penetrate values to eliminate those challenges.

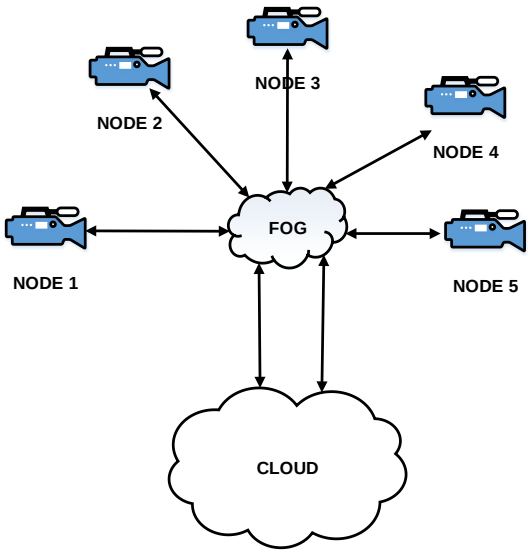


Figure 4.1: Fog based distributed video surveillance system

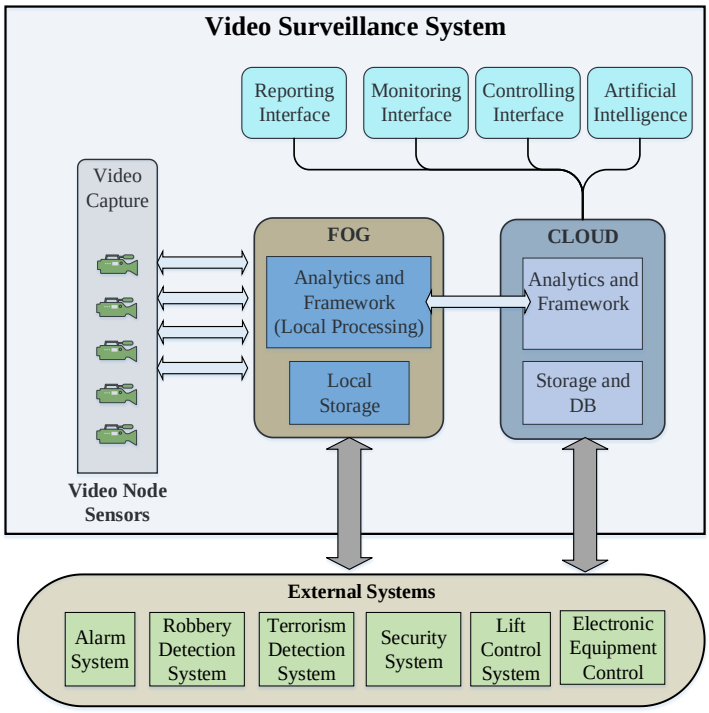


Figure 4.2: Fog-cloud collaborative services for video surveillance system

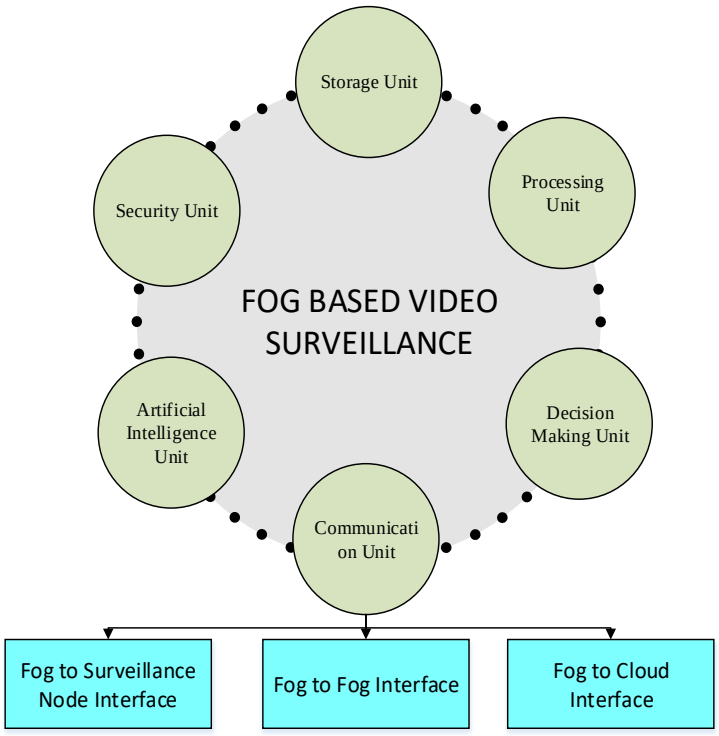


Figure 4.3: Multidisciplinary Units of Fog based surveillance system

4.2 Challenges in IoT based video surveillance

The emerging IoT platform introduces many new challenges in video surveillance system. Chiang et al. in [51], the authors summarize the opportunities and challenges of IoT in general, focusing primarily in the networking context of IoT. Here, several fundamental challenges in the context of video surveillance are discussed.

4.2.1 Latency Requirements

Video surveillance system often needs hard real-time operation. Crimes such as breaking doors or locks of a highly secured area need to be detected within a second or some millisecond. Automatic decision making (such as shutdown the lift) in a critical situation (such as stampede, fire) need to be executed immediately. These requirements fall far outside what mainstream cloud system can achieve. For example, in bank scenario, when a robbery is taking place, it should be detected with minimum latency. As the bank is a highly sensitive area, any latency of detecting this type of awful activity can cause a huge damage in respect of money and people's security. The early it can detect those activities, the easy would be to recover from the damage.

4.2.2 Network Bandwidth Constraint

A video surveillance node generates an exponential amount of data in every minute. Suppose, for simplicity, if a video surveillance node produces a video with 10 frames per seconds and frame resolution is 640x480 and there are thousands of nodes installed in a video surveillance system, it needs around 8.78 gigabytes/second of data transmitted to the cloud in an uncompressed way. This creates a huge bandwidth bottleneck in a communication system. For this reason, transmitting each and every video information to the cloud require prohibitively high network bandwidth. It is often unnecessary or sometimes prohibited due to regulations and video data privacy concerns. ABI research estimates that 90% of the data generated by the endpoints of a system will be stored and processed locally rather than in the cloud. For example, if the bank branch is the head office it should be occupied in a multistoried building which may need thousands of cameras. So if all the cameras send all the information to the cloud in 24/7, it would not be wise for the communication system and bandwidth. Because there are many times in the day for a camera where no activity can be detected. So it will be wise to control the data rate by introducing some intelligence.

4.2.3 Resource Constraint Video Surveillance Node

Modern video surveillance nodes have limited resource constraint include sensors, video buffer etc. From the computing perspective, these resource constraint video nodes will not be able to rely solely on their own limited resources for computationally complex work like object tracking, object recognition, face recognition etc. Interact directly with the cloud for all these requirements will be unrealistic and cost prohibitive as well. Because if the transmitted data is a secure video, then it requires resource intensive processing and complex protocols.

Modern video surveillance node has firmware integrated on the chip. Sometimes, the vendor of the node may need a firmware update on the fly. Each of these nodes which required heavy cryptographic operations and sophisticated procedures to obtain firmware updates from cloud services will be impractical. On the other hand, from a cloud perspective, updating same firmware for a number of nodes required same data transfer to all the nodes which is also impractical.

4.2.4 Time Critical Video Surveillance System

The uninterrupted and safe operation of video surveillance is often the top priority. Taking a system offline for any reason can miss sophisticated detection like theft or crime which may introduce significant security issue, business loss or intolerable customer inconvenience. In bank scenario, some surveillance cameras may be in very sensitive places. For example, a camera which monitoring bank vault needs to run without any intervention and any downtime can cause a big threat. Few other time critical cases are as follows:

- Requiring video nodes to be brought to vendor just to install firmware update packages can cause intolerable inconvenience and result in heavy cost and time consuming task for the system and vendors.
- Many automatic manufacturing systems based on video surveillance such as car assembly plants and electrical power generators in the energy grids need uninterrupted safe and correct operation and requires weeks to months lead times to plan for surveillance system shutdown.
- Many time critical control application which is based on video surveillance needs to be updated over time, cannot be moved to the cloud due to delay, bandwidth or other constraints. So it needs a new architecture for video surveillance system which is used in the mission-critical system.

4.2.5 Uninterrupted services Due to Intermittent Connectivity to The Cloud

Cloud services will have difficulty providing uninterrupted services to devices and systems that have intermittent network connectivity to the cloud. For example, in the rural areas, network connectivity may always not possible uninterruptedly. Resource constraint-based video surveillance node does not have such a large buffer to save all the frames in network downtime. Moreover, in time-critical video surveillance system, frame loss can be a big issue in downtime. Decision making upon an anomaly cannot be taken if the intelligence is installed only in the cloud. So it is impractical to install a time critical system in an environment where a decision can only be made from the cloud.

The cyber-physical system can be threatened by the intruder. If a video surveillance node is compromised by the intruder, it must be detected immediately without any time delay. In the interrupted network system, the cloud cannot detect in time that a node is being compromised. In

bank scenario, if the connection with the cloud cannot be established in a critical time like robbery, it would go towards a failure of detecting the robbery in real time. Moreover, if anyone tries to compromise the surveillance system of the bank due to any ill motive and the compromised situation can only be detected from the cloud, unstable connection with cloud also hamper detecting this situation. So it will also introduce security issue if video surveillance is in the traditional cloud system.

4.2.6 Security Challenges

Cloud-based security system works with today's internet which mainly designed for protecting enterprise network and data centers using firewalls. This system cannot be designed to protect a video sensor node or device from being compromised. So it needs new security paradigm for video surveillance system.

- Introduce Security Credential in Video Surveillance Node –

  Traditional video surveillance node has no security introduced on it. Traditionally a node connected in a system always sends the video to the receiver side. But in secured IoT-based video surveillance, the node should only send video data to authenticate receiver. On the other hand, Receiver has to sure that the video data are coming from authenticating Node (Not an intruder's data). So the video surveillance node needs its security mechanism in its firmware for the proposed IoT-based video surveillance system.
- Protecting Resource Constraint Video Surveillance Node –

  Replacing all the video surveillance nodes with IoT-based modern nodes sometimes is not appropriate for many reasons. Moreover, there may have some vendors which are not interested applying security credential in their video surveillance node. Therefore a fundamental question arises: How to protect a very large number of resource-constrained video surveillance node from security attacks.

  A large video surveillance system needs to run year after year. The video surveillance node may have very long lifespans and the hardware and firmware on them may be impractical to upgrade. Yet, these nodes need to remain secure over their long lifespans. Security threats will become significantly more advanced, many new threats will introduce and so the IoT-based video system mechanisms required to combat the growing threats will need to be enhanced and upgraded accordingly.

- Assessing the Security Status in Trustworthy Manner –

  A large video surveillance system will need to ensure the security of thousands or more video surveillance node. So it is essential to operate securely with a large number of video surveillance node. The situation which should be considered as follows:

  a) Video surveillance node can be installed in the physically unprotected environment can be compromised.
  b) Attackers can compromise a video surveillance node and damage the node while sending the background frame continuously as it appears normal.
  c) Video surveillance node may have security mechanism integrated into it, but if only the video sensor is compromised then it may not be detected by the security mechanism. So compromising any sensor of the video surveillance node need to be detected immediately.

- Responding to Security Compromises –

  Traditional responses after compromises rely on brute-force mechanism such as shutting down the system, reinstalling and replacing the components. But for mission-critical video surveillance system, it needs uninterrupted and safe operation even when it is compromised, if possible. For example, the bank's video surveillance system may be infected by a malware with the intention of theft the bank vault. Shutting down the video surveillance system can help the intruders to do this and cause significantly more damage including causes vast economic losses for the bank. Here are some other scenarios where interruption can be severe for video surveillance system.

  a) Think of an automatic robotic control manufacturing system depending on video surveillance. This type of system needs uninterrupted operation and safety over system integrity. This means that firmware update of surveillance node can only be installed in system scheduled down times which have to be short in time.
  b) Think of a drone for security surveillance flying midair is turned off because a security compromise of surveillance node is detected. Any security thread can remain undetected for the downtime of the surveillance.

Therefore, the conventional disruptive incident response system will no longer be adequate for mission-critical video surveillance system.

4.3 Advantages of Fog Computing in Video Surveillance

Computing, storage and networking resources should be shared by cloud and fog in video surveillance system. However, as fog stands at the edge of the network implies a number of advantages that make the fog a non-trivial extension of the cloud in this circumstances. The top-level view of the video surveillance system is shown in figure 3.4. Proposed video surveillance system has three layers. Sensor layer is the bottom layer which consists of the video surveillance nodes. The middle layer is the fog layer which receives the video frames from the sensor layer. Fog layer mainly responsible for many video analytics jobs includes meeting the real-time requirements for intelligent surveillance system described earlier. The upper layer which is cloud layer is responsible for long time archiving and off-loading the different analytic jobs. In reality, cloud and fog will share their responsibility according to the situation and requirements. The switching and routing devices will act as the fog device as future switches and routes should have the program interface integrated on those. Cisco has already developed an IOx framework of networked fog devices including router and switches so that customer can develop, manage and run a software application on those devices [19].

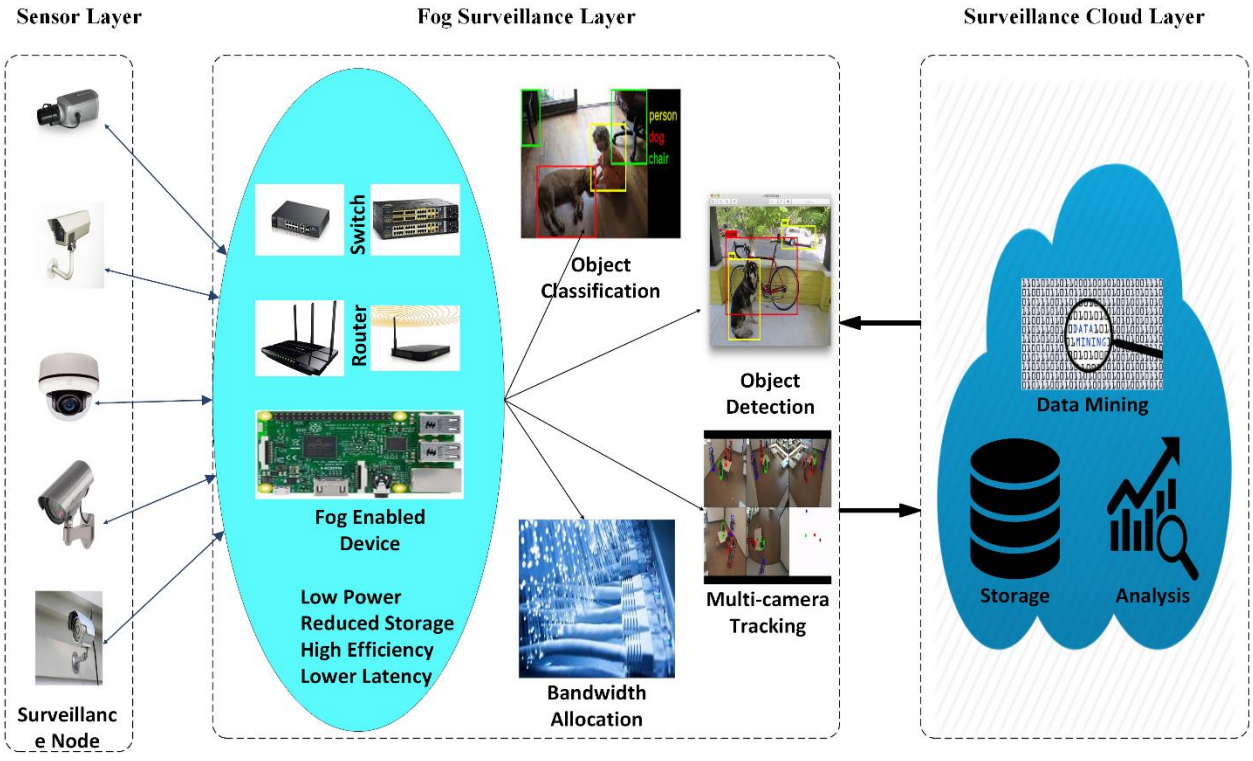


Figure 4.4: Layered architecture of the proposed IoT based video surveillance framework for power efficient, low latency and high throughput analysis of the surveillance big data

4.3.1 Low latency

The influence of the fog can be traced to support endpoints with rich services at the edge of the network, mainly for the applications with low latency requirements. So for the hard real-time video surveillance system, where low latency is must, best appropriate with fog. If the analytics tasks can be done at the edge of the network and an intelligent decision making can be done from the edge, it will be much faster to prevent any ill motive comparing to cloud-based system.

4.3.2 Network Bandwidth

In fog-cloud collaborative architecture, many functions are provided at the edge of the network which saves significant bandwidth. In cloud-based video surveillance system, each video frames have to travel to the cloud for further processing and archiving. In contrast, fog can process video frames, store frames and can make an intelligent decision at fog node which can avoid a significant number of frames transmitting to the cloud, and consequently saves bandwidth.

Some local processing is essential in many IoT-based applications as in video surveillance. Without fog, local processing is sometimes done by the resource constraint surveillance node which creates a barrier for the different vendors. Fog can eliminate this issue as local processing should be evaluated at the fog node.

4.3.3 Geographical Distribution

The services and applications targeted by the fog demand widely distributed deployments in contrast to the more centralized cloud [4]. Video surveillance system also spread in a distributed manner. For example, in bank scenario, there may have many branches spread all over the country. All the branches are under the same surveillance system. In cloud-based system, all the nodes of every branch usually directly connected with the cloud. But in fog based architecture, different fog can be run in different branches or geographical locations. All those fogs should be connected to the cloud at the end. Moreover, multiple fogs can be created in the same branch of the bank. Suppose if the branch is consist of five floors, five different fogs can be established which are not only connected to the cloud but also connected with each other. Figure 4.5 describes the connections in multi-fog based video surveillance system. In the multi-fog based system, there should be three communication interfaces for every fog for video surveillance system. Firstly, fog to node interface which provides the communication between camera nodes and fog in order to

transfer video frames and provide the required node settings. Secondly, fog to fog interface which provides the communications in an emergency situation or in a time of congested network situation. Thirdly, fog to cloud interface which provides all the communication between fog and cloud for collaborate archiving, processing and decision making. Different surveillance nodes can be distributed in different fogs according to their privileges, performance, and importance based on the requirement of the application.

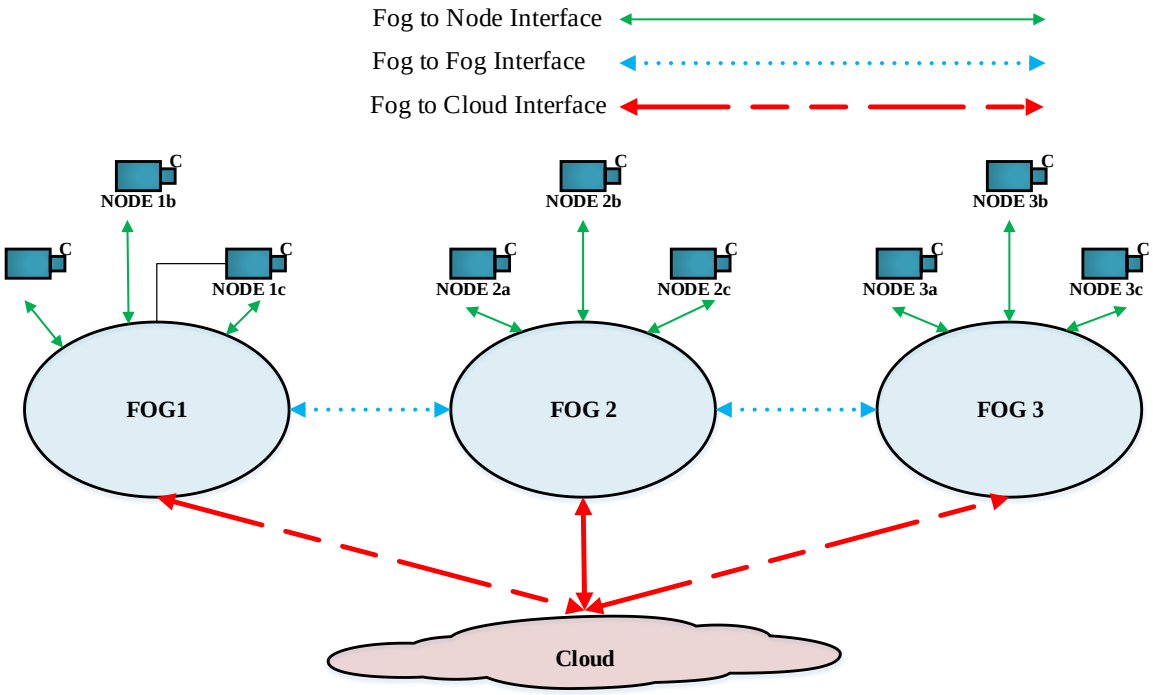


Figure 4.5: Multi Fog based distributed video surveillance system

4.3.4 Large Number of Nodes

Fog is suitable to implement in a large-scale sensor network coherently a distributed system which requires distributed computing and storage resources. Fog is also great wherever there are a large number of distributed nodes as a consequence of wide geo-distribution. In this scenario, distributed video surveillance system comply with the requirements as thousands and even more nodes can be connected in a large distributed video surveillance system.

4.3.5 Real-Time Interactions

Fog applications are suitable in real time interactions rather than off-line batch processing, certainly a good feature for video surveillance system. If all the processing and decision making come from the cloud, the system fails to perform when network goes offline. Fog absolutely eliminate this problem as it may operate the intelligent surveillance even in offline as it works at the edge of the network. Therefore, Fog can provide a significant contribution for the time critical hard real-time operation based on video surveillance.

4.3.6 Heterogeneity

Fog nodes come in different form factors and will be deployed in a wide variety of environments [4]. Video sensor camera may be provided by different vendors and the camera can be installed in a different environment of the same surveillance system. So the feature of heterogeneous characteristics of fog can create value in video surveillance applications.

4.3.7 Interoperability and Federation

Fog component usually is able to interoperate with each other to distribute the services across the domain. So it can help seamless support of certain application such as video streaming of video surveillance which requires the co-operation of different providers. Fog will be helpful for video surveillance system because next generation video surveillance system also intelligently communicate with some external systems.

4.3.8 Security

Fog, cloud and the surveillance node vendor - all will be responsible for the security credential on the surveillance node and the security system. Updated security mechanism can be integrated with firmware update of the surveillance node. Whenever it is detected that a video surveillance node is compromised, fog can provide the security mechanism instantly to overcome the latency issue. Fog can act as proxies for video surveillance nodes to help. Manage and update the security credential and firmware on these devices. It should monitor the security status of surveillance devices. It may detect threats in a timely manner by taking the advantages of local video information/security information and context to detect threats.

4.3.9 Scenario-Based Example

Now let consider a robbery scenario in a branch of the bank and describe how fog-cloud architecture reacts in this particular scenario. Figure 3.6 shows the sequence diagram of the scenario. For simplicity, only three surveillance cameras are shown in this scenario. Fog collects video frames from the cameras as usual. Fog always analyses with the statistics and machine learning. Suppose it detects a potential robbery from the frames of camera 1. It then gives a signal about this potential robbery to cloud and take some feedback from the cloud. Fog also calculates and subsequently decides that Camera 2 and 3 can give more information about this potential robbery (may be camera 2 and 3 are nearer to camera 1 or may they have significant relations with Camera 1). Fog then set high priority to those camera nodes and continuously taking frames from those with a high preference. After collecting more information from the multi-camera, fog can confirm whether it is actually a robbery or not. If fog is convinced, then it sends the information to the cloud and takes feedback from the cloud. Fog may provide some actions based on the learning and may fire an alarm, send SMS accordingly. Fog may then communicate with the building security system and robbery detection system according to the situation. After all the action completed and fixing the situation, fog then incrementally update itself by learning from the situation and update cloud about the whole scenario. Cloud also update itself and transfer the new knowledge (robbery scenario) to other fogs. The main advantages of this architecture are the real-time analysis and detection by the fog. If in this critical situation fog discovers that the cloud server is in offline, it should not wait for cloud feedback, rather than performs according to the situation with fog's internal intelligence.

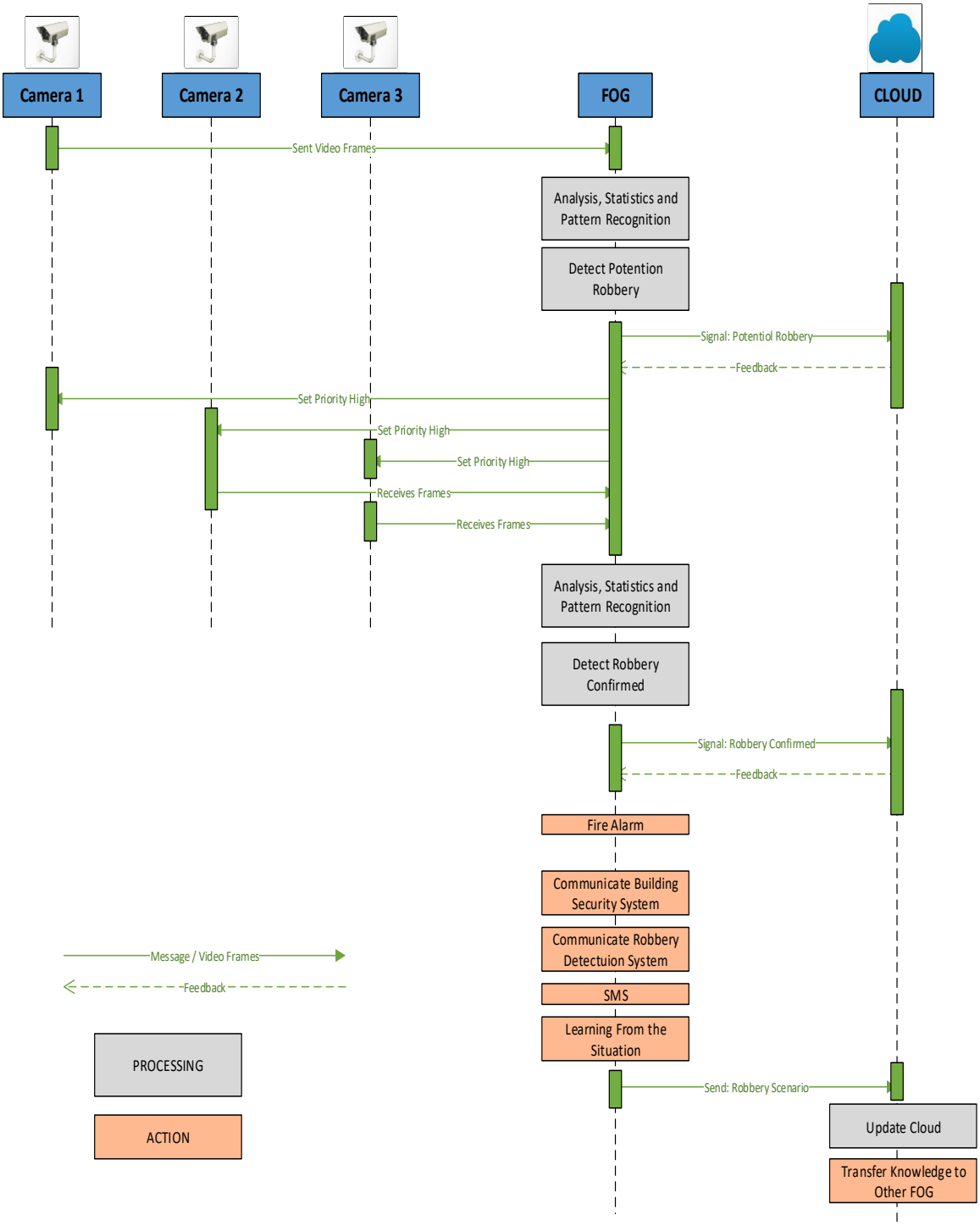


Figure 4.6: Sequence Diagram: Responsibility of video surveillance system in Fog-Cloud collaborative design towards a robbery detection scenario of a bank

4.4 Summary

In summary, this chapter established a conceptual framework for IoT based video surveillance system. This chapter also shows a bundle of challenges to integrate IoT in video surveillance system. Fog computing is introduced here to deal with those challenges and it is shown that Fog has enormous opportunities to eliminate those issues. At last, with a sequence diagram of a real-world scenario to realize and recognize the advantages of Fog considering intelligent video surveillance is shown.

Chapter 5 - Case Study 1: Frame Sampling Technique in Bandwidth Constraint IoT Environment

5.1 Overview

In this chapter, a power-efficient frame sampling technique that is applicable to the IoT and IoVT based video surveillance application is presented. The conventional way is to transmit all R, G and B components of all frames. Using this proposed technique, instead of sending all components, first one color frame is sent followed by a series of gray-scale frames. After a certain number of gray-scale frames, another color frame is sent followed by the same number of gray-scale frames. This process is repeated for video surveillance system. In the decoder, color information is formulated from the color frame and then used to colorize the gray-scale frames. To minimize the power and bandwidth consumption in IoT environment, this research includes a detailed analysis of proposed method using several color space converters, such as YCbCr, YUV, YEF, YCgCo and HSV.

This research presents a novel way that can save a significant amount of transmission energy without the need for additional hardware resources. Instead of transmitting all color images, the proposed IoT-based approach transmit only one color image followed by a series of gray-scale images. The gray-scale images received in decoder side are then colorized by a content-aware pre-processing and motion estimation and compensation technique. Conventional encoders such as H.264 or MPEG are computationally complex because they use motion compensation technique to remove temporal redundancy [99]. Since the energy consumption (in the form of computational cost) of motion compensation is very high and the video surveillance node has energy constraints, motion compensation is not integrated into video sensor node/encoder side. However, motion estimation technique is used in decoder side. The number of gray-scale images used can be controlled by the user or the system. Since no additional hardware is required in sensor node and gray-scale image sequence is mostly transmitted, a significant amount of transmission power can be saved, which results in an extended battery life of the surveillance camera.

5.2 System Design

The proposed technique aims to reduce IoT-based video transmission energy. Reducing the size of the transmission data with less reduction of quality of the video is the key target of this technique. The conceptual diagram of the conventional system and the proposed system are shown in figure 5.1.

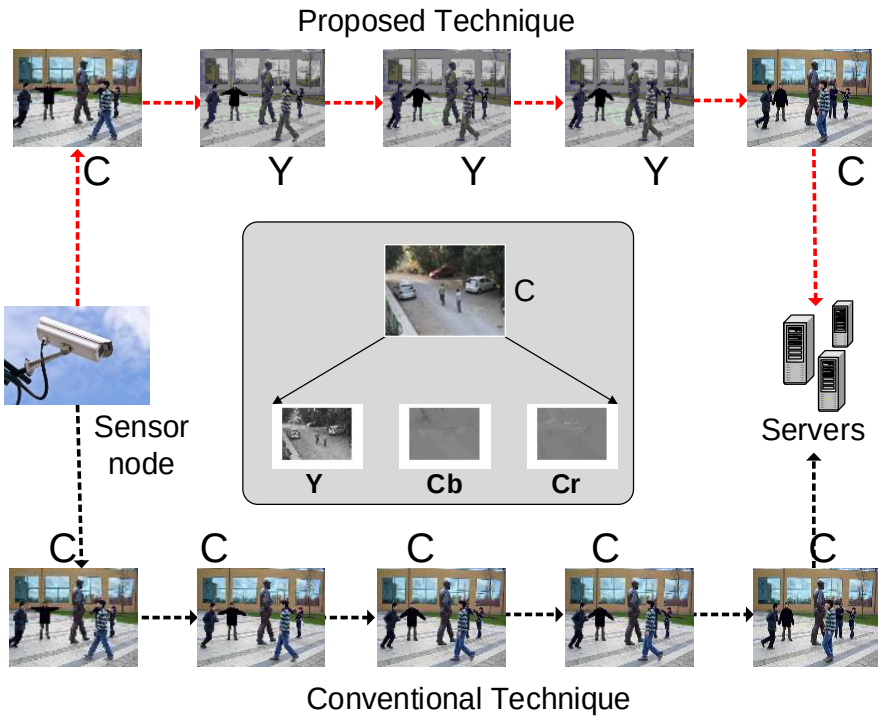


Figure 5.1: Proposed Vs Conventional Technique; Encoder sending a consecutive number of single component frames (Y) between three color frames in YCbCr color space.

The goal of this research is to design a system which can contribute to the following video surveillance constraints:

- As most of the IOT driven distributed video surveillance nodes are low powered and often battery driven, energy saving at those nodes are the key factor designing in IOT video surveillance application. Reducing transmission cost in encoder or transmission side can save significant energy. The proposed system targets to do most of the computation in decoder/receiver side. Decoders/receivers are mainly composed of distributed video servers where higher computation or higher energy is not an immense issue.

- In video surveillance real time application, changes within consecutive two frames are very low. So temporal information can contribute significant role for computational cost reduction.

In order to achieve the goals, the plan of this research is to scale down the size of the transmission data. To obtain this, for some images, the proposed encoder transmits the gray-scale component, not all three components. For remaining images, it transmits all the three components. In decoder, the received gray-scale images are colorized with proposed color image reproduction technique from the color frame which was received previously.

5.2.1 Encoder Design

The encoder of video surveillance node converts the frames to YCbCr color space. It then transmits one color image (3 component: Y, Cb and Cr) followed by a series of consecutive gray-scale (Y) images. The concept of encoder side is shown in Figure 5.2. For simplicity, let Energy Parameter (Ep) is considered as the number of gray-scale images used between two color images. Ep results on the bandwidth as well as required transmission power of a surveillance node. Higher Ep can save significant energy and bandwidth as it increases the number of gray-scale images while transmission. On the other hand, higher Ep has degradation effects on the quality of the reconstructed image.

Figure 5.2: Encoder sending a consecutive number of gray frames

This experiment does not let the encoder performing motion estimation in the encoder. However, it is already mentioned that the proposed system uses motion estimation in the decoder. A color video with YCbCr format, Y component which represents the luminance of the color, is a good candidate of motion estimation. Moreover, Cb and Cr components are normally in half

resolution of the Y component. The main advantage of selecting YCbCr color space for the proposed system is the correlation between Y and Cb, Cr which allows reproducing high quality of the color images with only Y component. Therefore, to reduce the transmission energy and to save computational cost at encoder side, only Y component is transmitted. Algorithm 1 provides the pseudocode for proposed IoT-based video surveillance encoder technique.

Algorithm 1 Encoder Technique

1: **for** each frame **do**
2: Convert the frame to *YCbCr*
3: Send all three components of *YCbCr* for every *Ep*-th frame, otherwise send only *Y* component.
4: **end for**

5.2.2 Decoder Design

The decoder is designed to maximize the quality of the images with less computational cost. The algorithm and the block diagram of the proposed system are shown in algorithm 2 and figure 5.3. When decoder receives an image with only Y component from the encoder, it sends the frame to a pre-processing step. In the pre-processing step, it calculates the temporal activity for each macroblock (MB) [78]. Though the standard MJPEG processes the image with the unit of 8x8 pixel block, different size of MB are used and the reconstructed frames are compared. Temporal activity for each MB is measured by frame difference (FD) with the previous frame. After each MBs activity is calculated, MBs with activity level lower or equal than the predefined threshold are marked as SKIP BLOCK. Color can be directly copied from the color of the previous frame for a SKIP BLOCK. MBs with activity levels higher than the predefined threshold are marked as MOTION BLOCK. That is, MBs that have more pixels with the significant difference from the previous frame are classified as MOTION BLOCK. For each MOTION BLOCK, the motion vector is measured from the previous frame. Motion estimation and compensation are used to colorize the MOTION BLOCK. Three step search block matching algorithm is used for motion estimation.

Algorithm 2 Decoder Technique

```
1:  for each incoming frame do
2:      if Frame is color then
3:          Save the frame
4:          Yp ← Y component of the frame
5:      else
6:          for each MB of the frame do
7:              amd ← absolute mean difference w.r.t Yp for corresponding MB
8:              if amd>predefined_threshold then
9:                  Use block matching algorithm for motion estimation to
                    reconstruct color frame
10:             else
11:                 Copy the color of previous frame to reconstruct current
                    frame
12:             end if
13:             Save reconstructed color frame
14:             Yp ← Y component of the reconstructed frame
15:         end for
16:     end if
17: end for
```

In order to analyze moving object detection performance of the proposed preprocessing, the results of frame difference operation are examined. MOTION BLOCKs of the image without the SKIP BLOCKs are shown in Figure 5.4 (b), higher number of MOTION BLOCKs need more processing time to reconstruct color image in decoder side. However, the activity level of MBs has been measured and found that a big number of MBs having very little or no activity in the surveillance video. In Figure 5.5, activity level of MOTION BLOCKs and SKIP BLOCKs are shown. If 15% of the maximum possible absolute difference is used as threshold for a video with low temporal activity, it is seen that about 98% of the MBs can be classified as SKIP BLOCK. As mentioned earlier, colors for the SKIP BLOCKs are directly copied from the previous frame's color, proposed technique can save significant computation energy for color reconstruction.

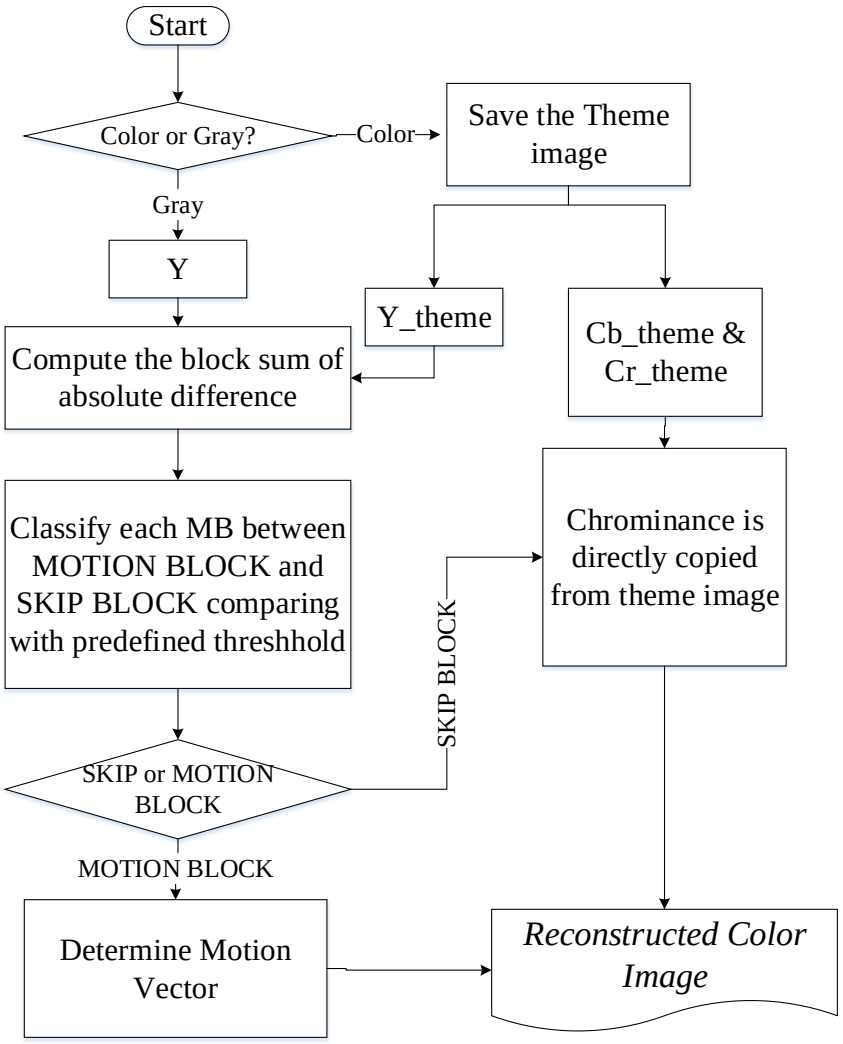


Figure 5.3: Flowchart for Decoder Technique

(a) (b)

Figure 5.4: Result of the proposed pre-processing operation. (a) Original frame image Y component. (b) MOTION BLOCK of the corresponding frame

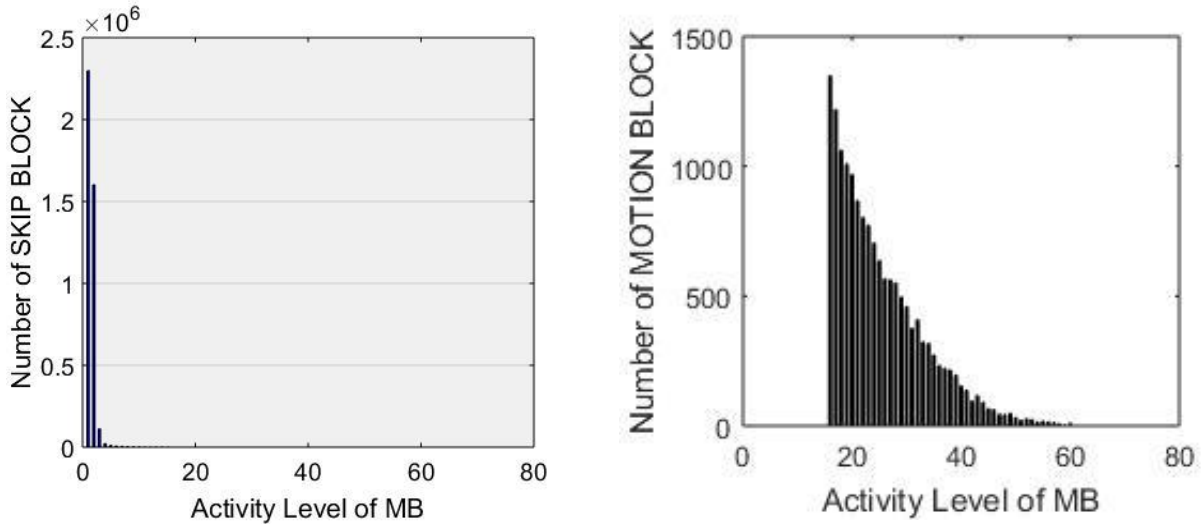


Figure 5.5: Result of the activity level of MB. (a) Activity level of SKIP BLOCKs (b) Activity level of MOTION BLOCKs

5.3 Experiment in Different Color Space

Not only in YCbCr color space, the experiments extended with proposed system experimenting with different color space. The extended experiment includes by transforming YCbCr, YUV, YCgCo, YEF [100] and HSV color space. All of those color space except HSV comprise the luminance (Y) and two color different component (e.g., E and F for YEF color space). For those color space, Y component is sent when transmitting single component. Equation 5.1, 5.2, 5.3 and 5.4 reveal the conversion matrices for converting from RGB to YCbCr, YUV, YCgCo and YEF respectively. YEF color space is designed the unique properties of endoscopic images for better compression [100]. From equation 5.3 and 5.4, it is seen that the conversion between RGB to YCgCo and YEF involve only a few addition and shift operation. So cost for converting from RGB to YCbCr is much higher than converting RGB to YEF or YCgCo. Component V is sent to HSV color space instead of Y of other color spaces. The equations from RGB to HSV conversion are given in equation 5.5, 5.6 and 5.7.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \tag{5.1}$$

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.288 & 0.436 \\ 0.615 & -0.515 & -0.1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{5.2}$$

$$\begin{bmatrix} Y \\ C_g \\ C_o \end{bmatrix} = \begin{bmatrix} 0.25 & 0.5 & 0.25 \\ -0.25 & 0.5 & -0.25 \\ 0.5 & 0 & -0.5 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{5.3}$$

$$\begin{bmatrix} Y \\ F \\ E \end{bmatrix} = \begin{bmatrix} 0.25 & 0.5 & 0.25 \\ -0.125 & 0.25 & -0.125 \\ 0.125 & 0.125 & -0.25 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} \tag{5.4}$$

$$H = \begin{cases} \left(\dfrac{G'-B'}{MAX-MIN}\right)/6, \text{ if } R'=MAX \\ \left(2+\dfrac{B'-R'}{MAX-MIN}\right)/6, \text{ if } G'=MAX \\ \left(4+\dfrac{R'-G'}{MAX-MIN}\right)/6, \text{ if } B'=MAX \end{cases} \tag{5.5}$$

$$S = \frac{MAX - MIN}{MAX} \tag{5.6}$$

5.4 Result and Analysis

For performance comparison of the proposed system, baseline MJPEG with variable distortion and compression is used. The quality of the decoded video is measured by software models based on MATLAB. Five surveillance video sequences with varying temporal activity are used, available at [101].

5.4.1 Savings in Transmission Energy

Energy consumption of the encoder side needs to keep low in order to be successful in the integration of an IoT sensor node. It is noted that the color reproduction is implemented in receiver/server and no additional hardware is needed in the video surveillance node to implement it. The power consumption in terms of transmission power is computed for the conventional (all three frames) and proposed IoT-based cases. Here, compression Ratio of the no compression,

lossless compression and lossy compression are 0, 60 and 80% respectively [77]. Energy consumption per bit of the sensor node measured to be 21.05 nJ/bit [77]. Figure 4.6 shows the average transmission energy consumption per frame for a video of size 640 x 480. It is seen, for conventional transmission (all frames are color), the energy consumption is the highest for all cases: 165.98 mJ for uncompressed, 66.4 mJ for lossless and 33.2 for lossy images. Whereas using the proposed IoT based technique, the energy can be significantly reduced for all cases. For example, using Ep = 4 (one color frame followed by four gray-scale frames), the consumption is reduced to 78, 31.2 and 15.6 mJ for uncompressed, lossless and lossy cases respectively, which results in an energy saving of 53%.

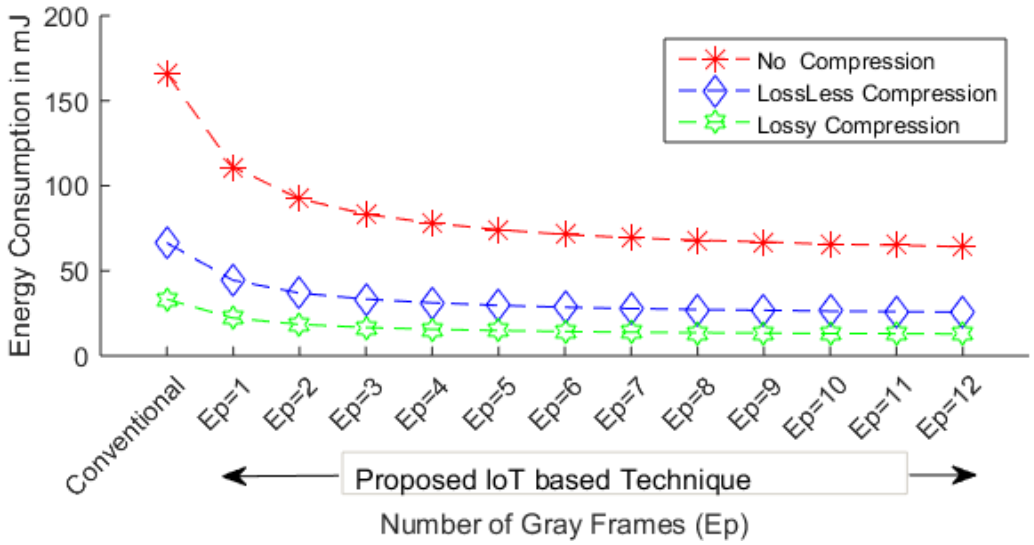


Figure 5.6: Savings in transmission energy consumption for different Ep in YCbCr

5.4.2 Bandwidth Savings

In Figure 5.7, estimated bandwidth savings for the different number of Ep is shown. It is seen that higher Ep introduces noise for reconstructed videos. But increases Ep after certain level can not improve the bandwidth savings. With this finding, this case study concludes that sending 4 or 5 number of gray-scale images (Ep = 4 or 5) between two color images is a good choice in terms of energy consumption and quality of the video.

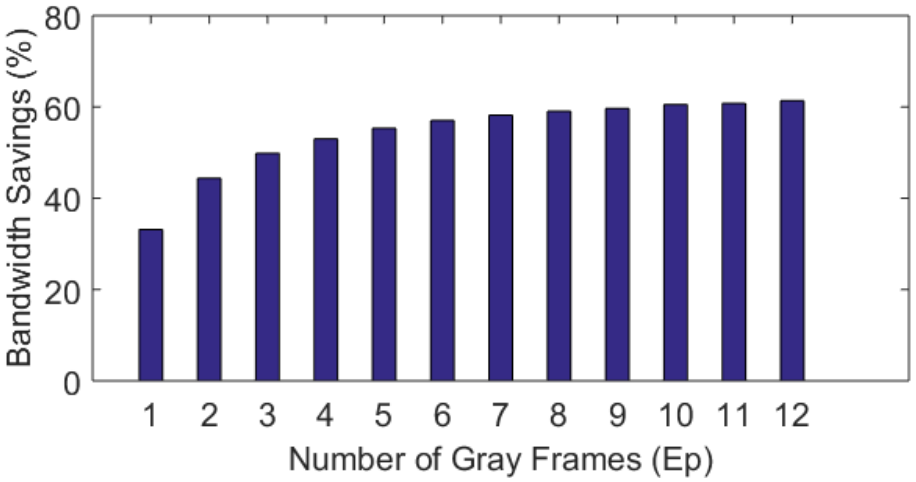


Figure 5.7 Performance Comparison of required transmission bandwidth for different numbers of Ep in YCbCr.

5.4.3 Video Quality for different Color Space

The experiment is done varying the number of transmitted consecutive single component frames (Ep) form 1 to 10 and computed peak signal to noise ratio (PSNR) and mean square error (MSE) of the reconstructed frames. It is found by experiment that in YCbCr, YUV, YEF and YCgCo color space average PSNR (dB) are much better than HSV color space. For example, for Ep=3, average PSNR are 46.24, 45.9, 45.96 and 45.8 dB for YCbCr, YUV, YEF and YCgCo respectively. But for HSV, it is much lower than others (43.33 dB). Experimental MSE values also reveal the similar result found for PSNR. For example, average MSE (for Ep=3) values are 1.73, 1.86, 1.81, 1.86 and 3.38 for YCbCr, YUV, YEF, YCgCo and HSV color space. The experimental result shows that proposed approach gives marginally better result in YCbCr color space for every combination. Figure 5.8 and Figure 5.9 Show PSNR and MSE comparison respectively for Ep = 1 to 5.

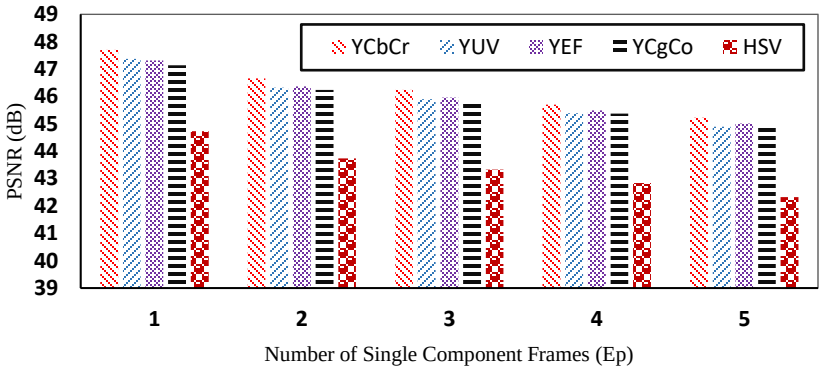


Figure 5.8: PSNR (dB) comparison for different color space

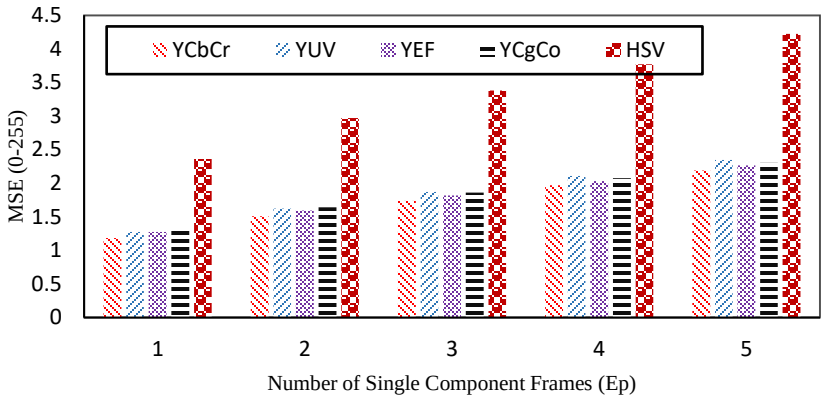
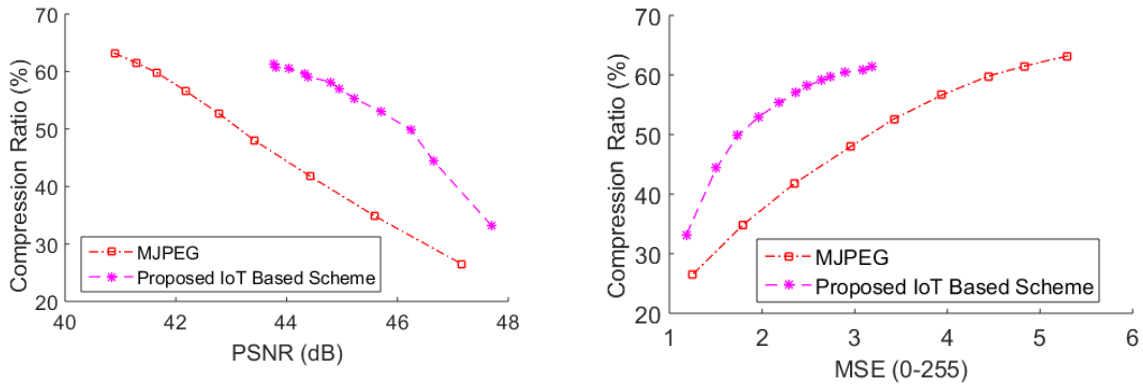


Figure 5.9: MSE (0-255) comparison for different color space

5.4.4 Rate Quality Characteristics

The proposed scheme not only reduces significant transmission energy but also it produces a high quality of videos as it is seen from rate-quality characteristics. In this technique, experimenting with different number of Ep, found different compression ratio and MSE/PSNR for that ratio. To compare, MJPEG is used as a standard for the same videos those examined for IoT-based technique. The experiment varies the quality factor of the MJPEG and produce multiple quality videos from the same video. Then compute the compression ratio, MSE and PSNR of those MGPEGs. MGPEG is a codec and so it comprises compression and sub-sampling inside the codec. But the proposed methodology is based on only component skipping technique without using any other compression or subsampling. It shows that even using any of those traditional compression techniques, proposed approach still give a better result for the same amount of compression.

The experiment varies the quality factor of the MJPEG and produce multiple quality videos from the same video. Compression ratio, MSE and PSNR of those MGPEGs are then computed. In Figure 5.10, it is seen that for a fixed compression ratio, proposed technique provides better video quality. For example, with 50% compression, computed PSNRs are 46 and 43.5 dB for proposed approach and MGPEG respectively. For the same compression (50%), computed MSE are 1.8 for proposed scheme and 2.9 for MJPEG. On the other hand, for a target PSNR of 46 dB, proposed scheme can compress 55% which is much higher than MJPEG (37%).



(a) (b)

Figure 5.10: Rate-quality characteristics of the proposed technique in YCbCr

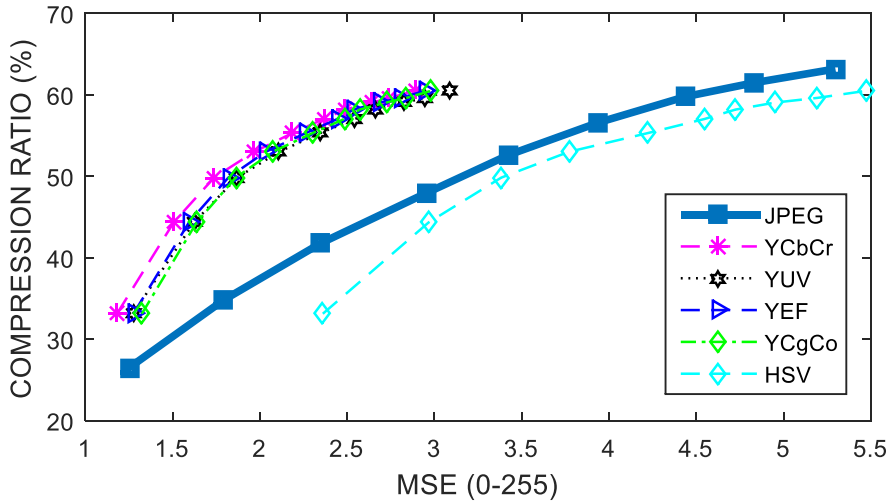


Figure 5.11: Rate-quality characteristics of the proposed technique and MJPEG for different color spaces

While experimenting in other color space, for example, for 50% compression, around 46 dB PSNR is found for YCbCr, YEF, YUV and YCgCo whereas around 43 dB is found for conventional MGPEG. For the same compression of 50%, MSE of around 1.9 is found whereas, for conventional MJPEG, it is around 3. On the other hand, for a target MSE of 1.8 proposed approach in YCbCr, YUV, YCgCo and YEF can achieve around 47% compression ratio, whereas, for the same MSE, conventional MJPEG can achieve only 35% compression. Figure 5.11 shows the curves of MSE vs compression ratio for different color space and conventional MGPEG.

5.5 Findings

The quality of the video in terms of MSE and PSNR of proposed system depends on the following:

1) Ep or number of consecutive gray frames.

2) The threshold to categorize SKIP BLOCK and MOTION BLOCK.

3) Macro block size.

Changing these parameters can contribute towards the whole video surveillance quality. By implementing proposed technique, the video surveillance application can dynamically change Ep and threshold (for classifying MB) with the bandwidth or required quality of the video. When a quality video is important than the transmission energy savings, the system can decrease Ep and decrease the threshold to reconstruct a high-quality video. On the other hand, when the system bandwidth or channel rate becomes low or operates under a constraint energy mode, it can increase Ep to save significant bandwidth.

The most important advantage of the proposed scheme is the high efficiency in terms of energy, bandwidth and computational cost. As there is no operation introduced other than the segmenting color component in the encoder, it is of low complexity. It is also bandwidth efficient because sending single component of a color image requires less bandwidth.

5.6 Summary

In summary, this chapter presented energy efficient color reproduction technique by frame sampling that is targeted for IoT-based video surveillance application. The proposed IoT-based design achieves low-energy consumption as well as good quality reconstructed video. This scheme

can enable sensor nodes in IoT to perform more operations with energy constraints. Therefore, the proposed design can be a lightweight preference to the application of video surveillance for IoT-based energy constraint sensors.

Chapter 6 - Case Study 2: Frame Reproduction Technique by Variable Length Pixel Encoding

6.1 Overview

In this chapter, a bandwidth efficient and low complexity frame reproduction technique that is applicable in IoT based video surveillance application is presented. Using this proposed technique, only the pixel intensity that differs heavily comparing to previous frame's corresponding pixel, is sent. If the pixel intensity is similar comparing to previous frame, the information is not transferred. With this objective, the bit stream is created for every frame with predefined protocol. In cloud side, the frame information can be reproduced by implementing the reverse protocol from the bit stream. This research is also extended by implementing Golomb-Rice encoding to make the system more bandwidth efficient.

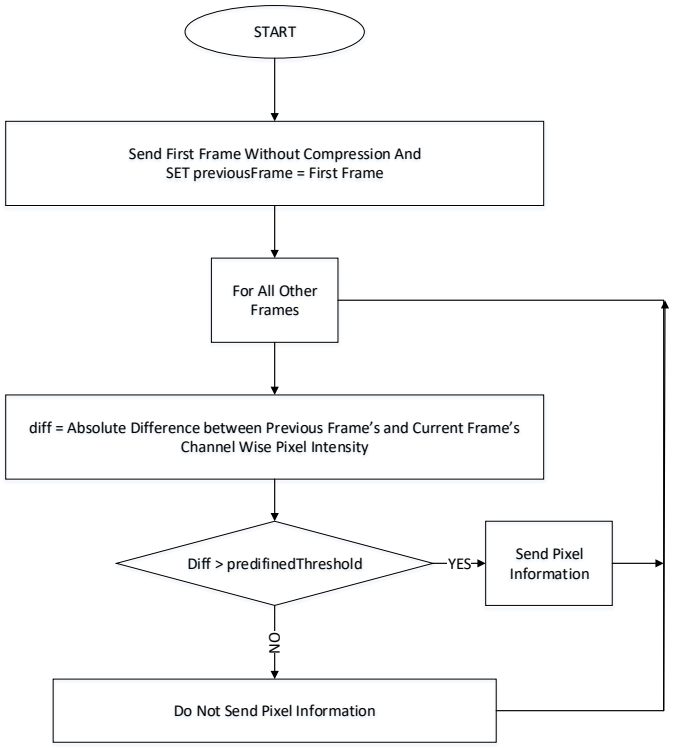


Figure 6.1: Flowchart of the proposed variable length pixel encoding

This research was executed in a couple of ways. Firstly, implement it RGB color frames, which can be implemented in Fog. Secondly, bayer images which can be implemented in

surveillance camera sensor to make the Fog light weighted. This research presents a unique way that can save a significant amount of energy without the need for additional hardware resources. Since no additional hardware is required and a significant amount of power can be saved, results in an extended battery life of the surveillance camera.

6.2 System Design

The proposed technique is designed in such a way that it reduces the transmission of redundant information of the previous frame. The first image of the video is transmitted in an uncompressed way from the Fog. For the subsequent frame, Fog compares the pixel intensity with the previous frame and send the corresponding frame information if required (if higher than a predefined threshold).

The cloud saves the first uncompressed image sent from Fog. For the subsequent images, cloud receives the bitstream and decodes the bitstream. If cloud finds pixel information in the bitstream, it saves the information. Otherwise, if it does not contain the information, the cloud takes the intensity from the previous image. Then it saves the reconstructed image for next image reconstruction. The block diagram of the proposed variable length pixel encoding is shown in figure 6.2.

In encoder or Fog, when the current image is saved for comparing the next image, it is not saved as the source frame or input frame. It is saved exactly the information that server/cloud will be extracted/decompressed from the bit stream. When comparing pixel with the previous image pixel, if it is compared with the different value in Fog and Cloud side, the error will be propagating for the next consecutive frames. For this reason, while comparing, always compare in Fog with the value which will be reconstructed in cloud side.

Figure 6.2: Block diagram of the proposed variable length pixel encoding

6.2.1 Creation of Bit Stream

The flow chart for adding a pixel information into bit stream is shown in figure 6.3. The steps of bitstream creation for an image are as follows:

1) Define three different thresholds for three color channels. The objective is to send the channel intensity to cloud if the absolute difference of the intensity between the current frame and the previous frame is greater than this predefined threshold. This threshold value determines the compression or bandwidth savings and quality of the image. If the threshold is large, that means it requires fewer pixels information to be sent. The information which is not sent is set as the information of the previous image.

2) The three different threshold value determines which channel information to be sent. It is usual for an RGB pixel to send three channel, two channel or single channel information depending on the channel intensity difference with the previous image's corresponding pixel.

3) For every pixel, the proposed system sends at least two bits, which corresponding the number of the channel to be sent. Depending on these two bits next subsequent bits are selected as follows.

a. If the two bits value is 0 (bit value 00), that means no information of this pixel will be sent.

b. If the two bits value is 1 or 2 (bit value 01 or 10), that means 1 or 2 channel(s) intensity will be sent next. Then it sends 10 bits for each channels information. Within these 10 bits, 2 bits are for channel identification (00 for RED, 01 for GREEN and 10 for BLUE) and the next 8 bits are for channel intensity.

c. If the two bits value is 3 (bit value 11), then all three channels (R, G and B) intensity will be sent one by one maintaining defined seriality. As all the three channels information has to be sent, no channel identification bits are required.

4) In this way, all the pixels information will be sent by raster scan (row major).

Figure 6.4 shows an example of the bit stream creation from the image. For simplicity, for example, the image has 4 pixels. The previous and current pixel's intensity is shown in the figure. All the three threshold values are 5 for this example. For pixel 1, the absolute differences between current and previous pixel are 1, 8 and 3 for R, G and B accordingly. Only the Green intensity differs much than the threshold, so this system only sends green intensity value. For this, it first sends 1 in 2 bits (Single channel's intensity), then channel identification 1 in 2 bits (Green) and 80 in 8 bits (Green intensity value). For pixel 2, all the absolute differences are less than the threshold, so no channel information will be in the bitstream. So only the number of channels 00 will be sent for this pixel. For pixel 3, all the channels intensities differ more than the threshold values. So it first send 11 (3 – 3 channels), then sends R, G and B intensity one by one.

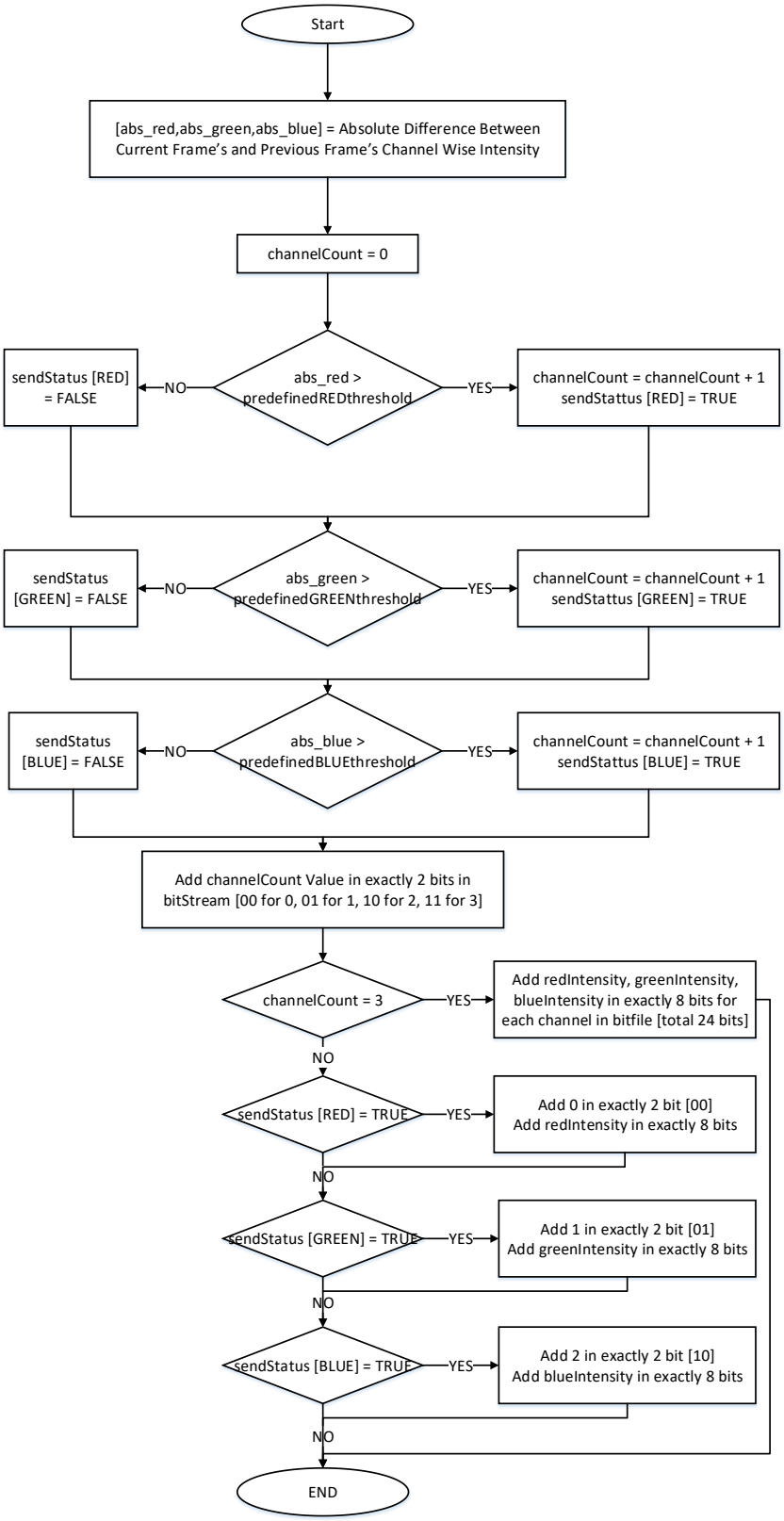


Figure 6.3: Flowchart for adding a pixel information into bitstream

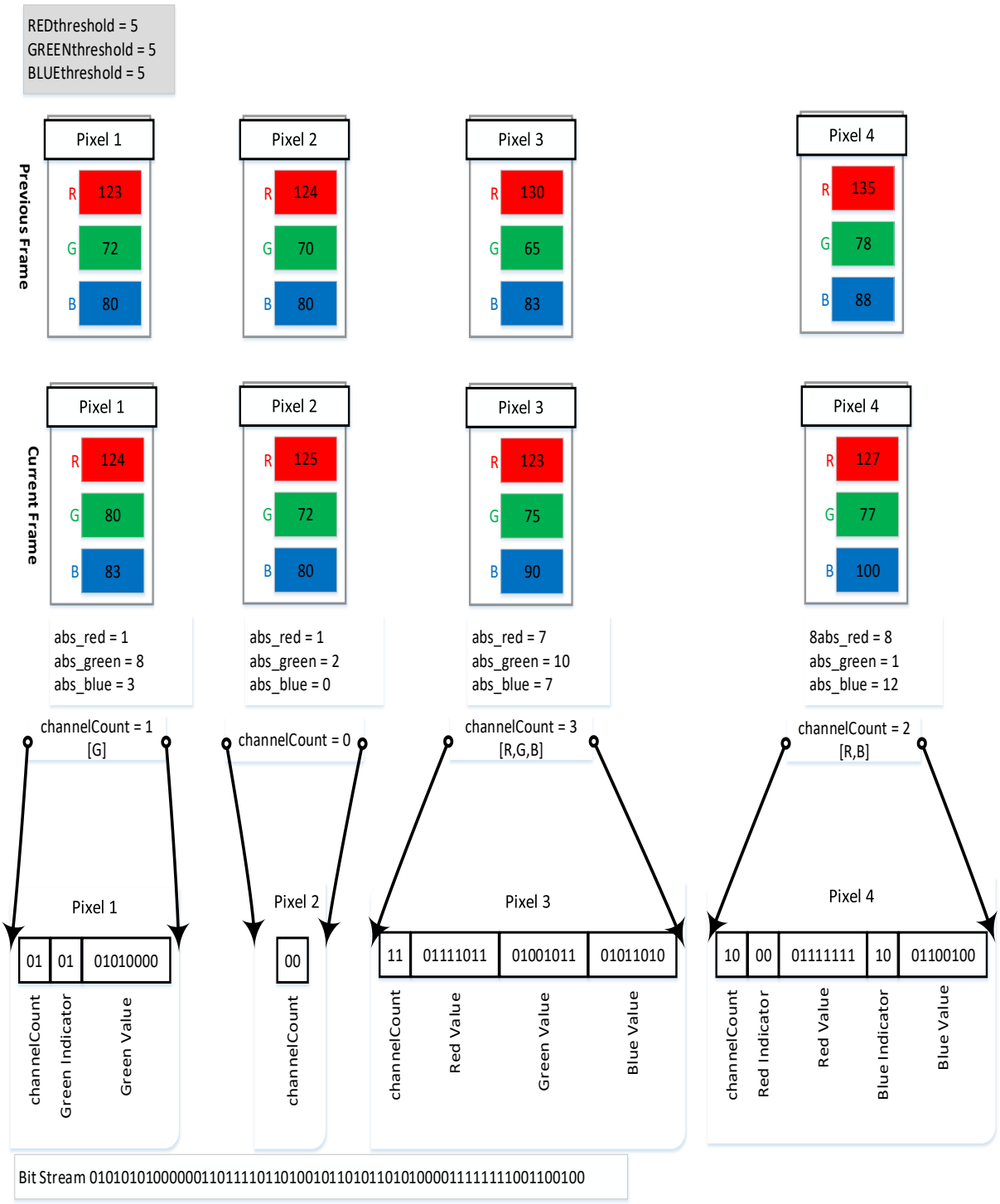


Figure 6.4: Example of bitstream creation from an image

6.2.2 Extract Pixel Information in Cloud Side

Extracting the pixel information is a reverse process. In cloud, for every pixel, 2 bits are read and determines how many channel information is received for that corresponding pixel. The extracting steps are as follows:

1) If first two bits are 00, that means no channel information is available in the bitstream for this pixel. The values of the pixel are same as the previous pixel.

2) If first two bits are 01 or 10, then it read next two bits and determine the channel indicator and read another 8 bit for the corresponding channel intensity.

3) If first two bits are 11, then it read next 24 bits and set the R, G and B values for that pixel.

The flow chart for extracting the pixel information is shown in figure 6.5.

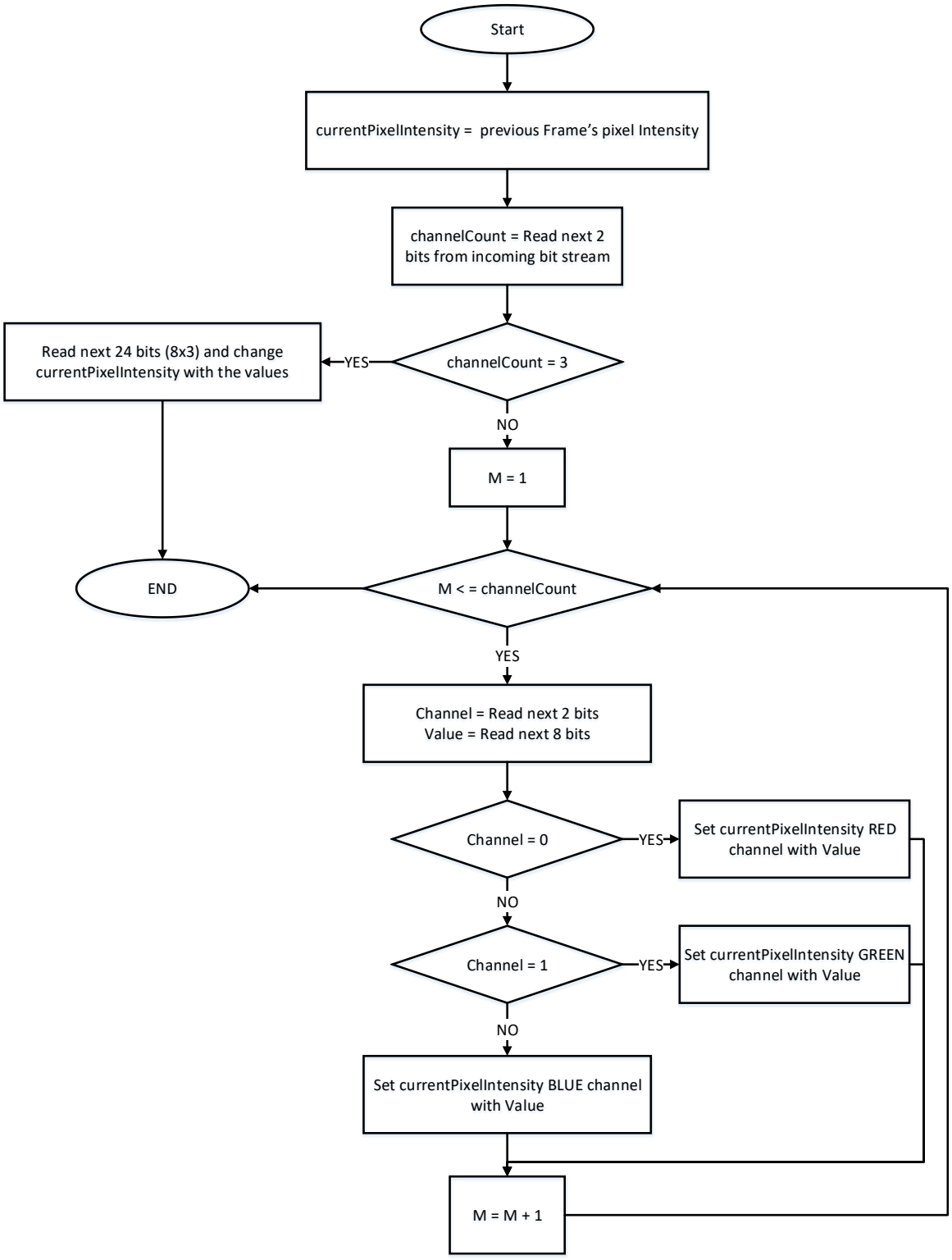


Figure 6.5: Flowchart for extracting pixel information from bitstream

6.3 Experiment for Bayer (RGGB) Video Image

This research also extends to experiment in Bayer image format. The Bayer format is found from manual transform from RGB image. The Bayer format is important in IoT, because it can add compression in node side. As Bayer video has only one channel for every pixel, the algorithm and format of the bit stream have to be changed. Differing from the process for RGB format, it needs only one threshold for comparing. Depending on that threshold, the Node determines whether it sends the pixel information or not. It takes 1 bit for this decision instead of 2 bits (0 means do not send, 1 means send the information). If it is 1, then another 8 bits for the intensity value is sent. Figure 5.6 shows the flowchart for adding pixel information into bitstream for Bayer images.

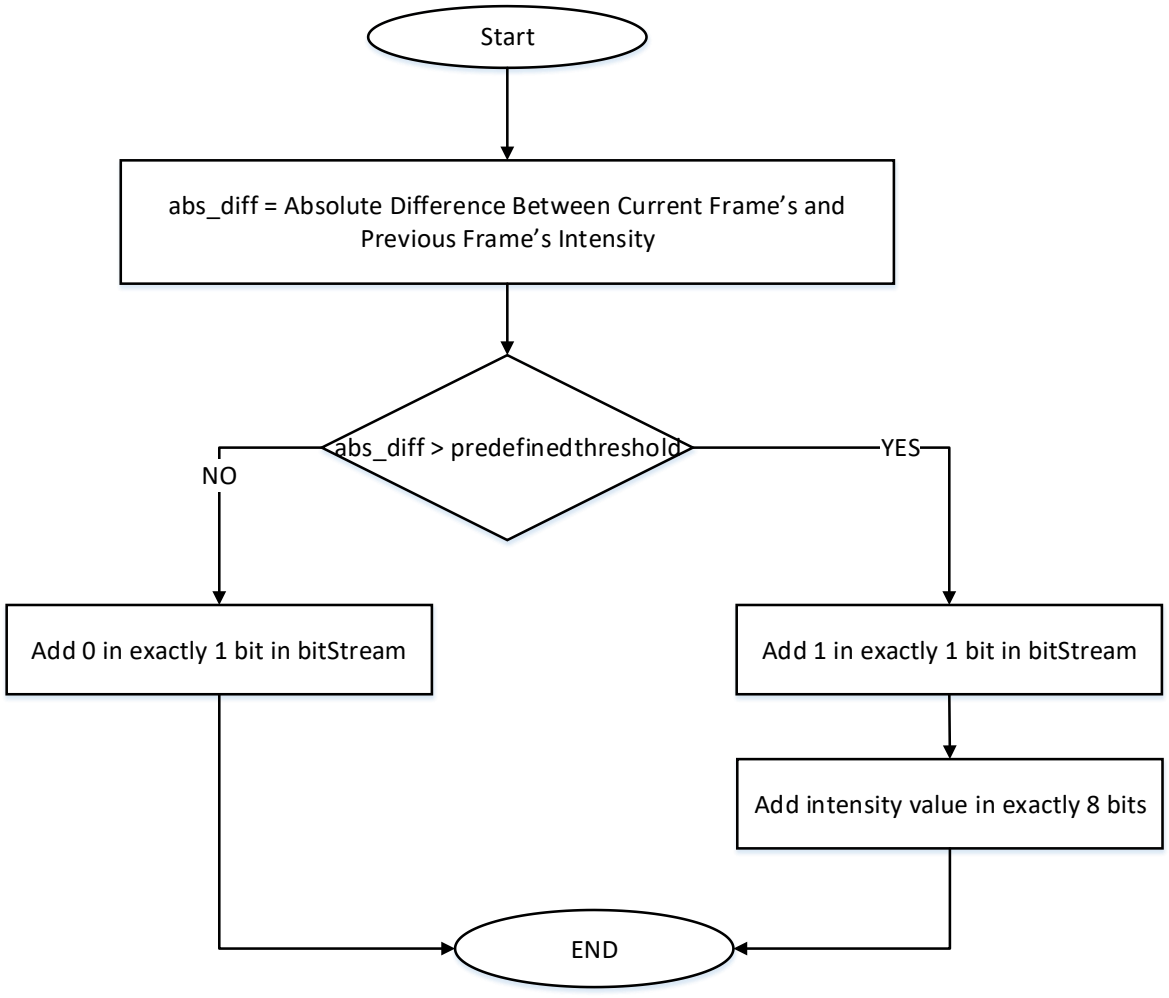


Figure 6.6: Flowchart for adding a pixel information for Bayer image pixel

6.4 Result and Analysis

For performance comparison of the proposed system, baseline MJPEG with variable distortion and compression is used. The quality of the decoded video is measured by software models based on MATLAB. We use five surveillance video sequences with varying temporal activity, available at [101] as Case Study 1.

6.4.1 Video quality

The experiments are done varying the threshold value. At first, the threshold value taken with big gaps to understand the quality in the different region of threshold. The threshold value for this experiment starts from 2 to 30 for R, G and B threshold. The average PSNR (in dB) and MSE (0-255) are calculated for video images. The PSNR found for threshold values 2 and 30 are 47.5 dB and 29.8 dB respectively. It is seen than PSNR value range is 40 to 47 when threshold value varies from 2 to 5. So this experiment further extends by experimenting with the threshold value from 0 to 5. For example, for threshold value 3, PSNR and MSE are 44.3 dB and 2.4 respectively. Figures 6.7 and 6.8 show PSNR and MSE respectively for different threshold values.

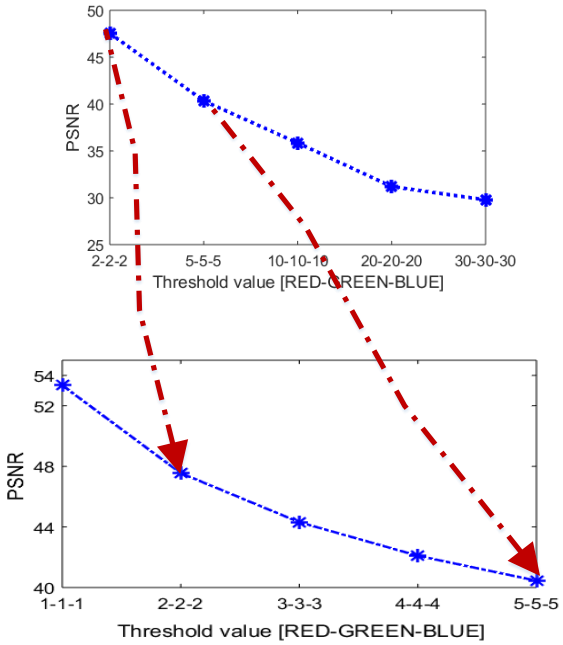


Figure 6.7: PSNR(dB) for proposed technique with RGB images

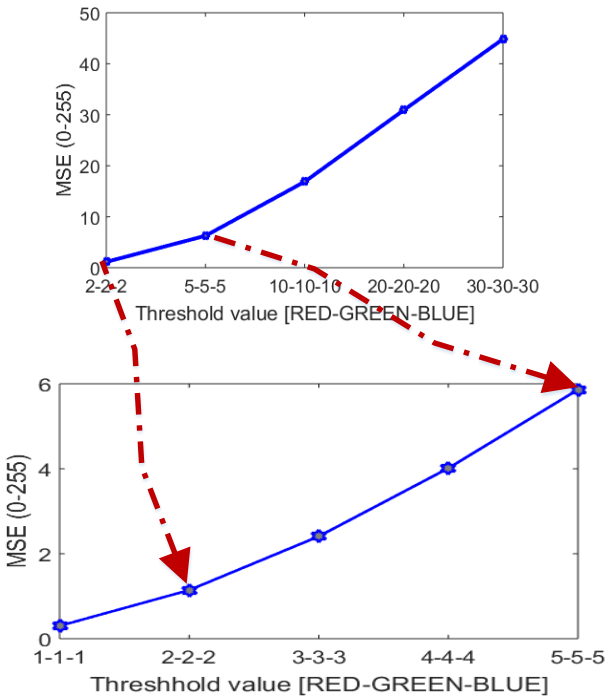
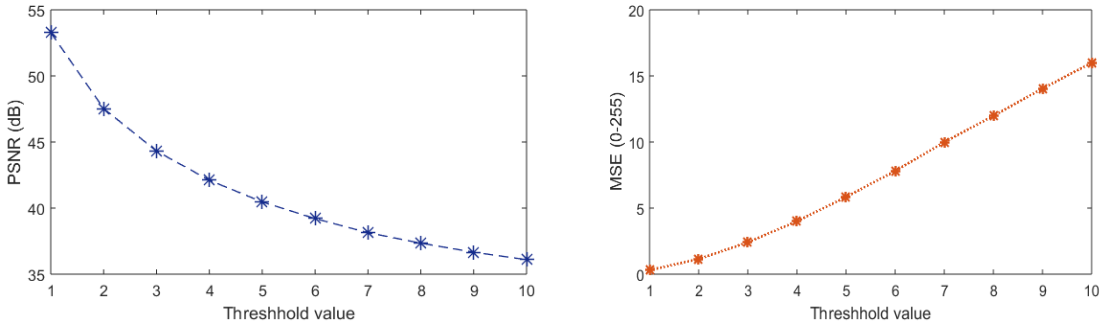


Figure 6.8: MSE (0-255) for proposed technique with RGB images

The experiments are done varying the threshold value for Bayer images. Changing the threshold for 1 to 10, the experiment calculates PSNR and MSE for Bayer images. For example, The PSNR and MSE found for threshold values 5 are 40 dB and 6.5 respectively. Figure 6.11 shows the PSNR and MSE values for the different threshold.



(a) (b)

Figure 6.9: PSNR (dB) and MSE (0-255) for Bayer images

6.4.2 Compression

In this proposed technique, if no information of the pixel is aggregated in bitstream, it too needs two bits for channel count (00). But in traditional way, it is needed to send 24 bits. When single channel intensity of a pixel in bitstream, it needs 12 bits and for two channels, it needs 22 bits of information in bitstream. If all the three channel intensities need to be send, it needs 26 bits of information. Therefore, percentages of bit saves are 91.66%, 50% and 8.33% for 0,1 and 2 channels respectively. But if it is needed to send all the three channels of a pixel, no bits can be saved as it needs extra two bits in bitstream. Table 6.1 shows the bit save information for different number channels of a pixel.

Table 6.2 shows channel counts in bitstream for a video and corresponding bit/pixel. It is seen that when threshold value increases, sending a number of information is less as it needs to send minimum pixel information. Suppose for threshold value 4, 78.7% pixels information need not be integrated into bitstream. These value can be copied from the previous frame. A number of pixels with single and double channels are 10.47% and 5.01% respectively. It is needed to send all the three channels intensity for only 5.45% of pixels. The needed bits per pixel for this video with threshold value 4 is 5.45 and 5.29 without and with implementing Golomb-Rice encoding respectively.

Table 6.1: Bit save for different number of channel in bitstream

Number Of Channels	Information	Total Bits needed	Total Bits in RGB	Bits save	Bits save (%)
0	Channel Count [2 bits]	2	24	22	91.66
1	Channel Count [2 bits] + Channel identification [2 Bits] + Intensity Value [8 bits]	12	24	12	50
2	Channel Count [2 bits] + (Channel identification [2 Bits] + Intensity Value [8 bits]) x 2	22	24	2	8.33
3	Channel Count [2 bits] + (Intensity Value [8 bits] x 3)	26	24	Need 2 extra bits	Cannot save bits

Table 6.2: Number of channels in bitstream and needed bits per pixel

Threshold [RED,GREEN,BLUE]	Number of pixels (%)				bits/pixel	
	Channel Count				Before Golomb-Rice Encoding	After Golomb-Rice Encoding
	0	1	2	3		
0 , 0 , 0	28.47	2.49	14.25	54.78	18.25	17.16
1 , 1 , 1	40.78	19.84	12.23	27.16	12.95	12.29
2 , 2 , 2	58.23	17.38	9.44	14.95	9.21	8.81
3 , 3 , 3	70.27	13.82	6.99	8.92	6.92	6.67
4 , 4 , 4	78.70	10.47	5.01	5.82	5.45	5.29
5 , 5 , 5	84.53	7.87	3.48	4.12	4.47	4.36
6 , 6 , 6	88.61	5.83	2.35	3.20	3.82	3.75
7 , 7 , 7	91.41	4.50	1.68	2.41	3.36	3.31
8 , 8 , 8	93.14	3.67	1.30	1.89	3.08	3.04
9 , 9 , 9	94.56	2.83	1.02	1.58	2.87	2.83
10 , 10 , 10	95.59	2.21	0.82	1.38	2.72	2.69

Table 6.3: Bandwidth savings and required bits/pixel for Bayer images

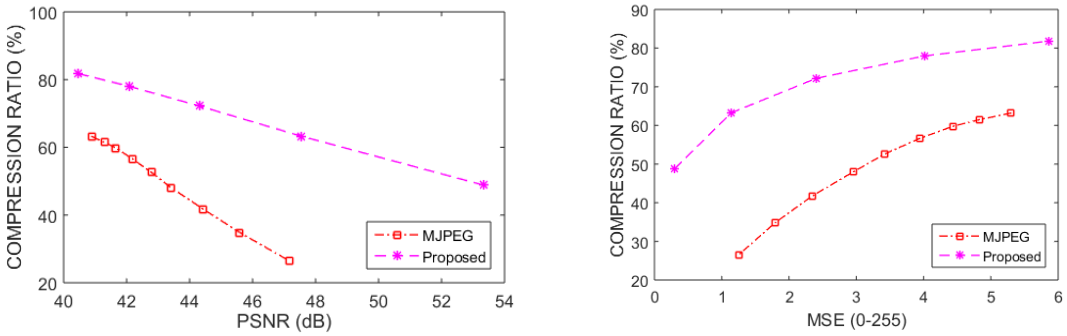
Threshold	Bandwidth savings (%)	Bits/pixel
0	75.22	5.95
1	82.77	4.13
2	86.90	3.14
3	90.18	2.36
4	91.90	1.94
5	93.01	1.68
6	93.74	1.50
7	94.26	1.38
8	94.59	1.30
9	94.83	1.24
10	95.00	1.20

Table 6.3 shows bandwidth savings and bits/pixel needed for Bayer images. It is seen that when threshold value increases, sending number of information is less as it needs to send minimum

pixel information. Suppose for threshold value 4, 91.9% bandwidth can be saved for Bayer images. The corresponding bits needed per pixel is 1.94 which is much lower for Bayer images. It is seen that from the table that even for lossless compression, it needs only 5.95 bits/pixel which is very significant.

6.4.3 Rate quality characteristics

The proposed technique not only reduces significant transmission energy but also it produces a high quality of videos as it is seen from rate-quality characteristics. In this technique, experimenting with a different threshold value, found different compression ratio and MSE/PSNR for that ratio. To compare, this experiment uses MJPEG as a standard for the same videos examined for proposed IoT-based technique. To compare with the conventional technique, MJPEG is used as a standard for the same videos examined for IoVT-based technique. The quality factors of the MJPEG are varied and multiple quality videos are produced from the same video. The compression ratio, MSE and PSNR of those MGPEGs are computed. MGPEG is a codec and so it comprises compression and sub-sampling inside the codec. But the proposed methodology is based on only variable length encoding and bit generation technique without using any other compression or subsampling. This experiment shows that even without using any of those traditional compression techniques, proposed approach still give a better result for the same amount of compression.



(a) (b)

Figure 6.10: Rate-quality characteristics of the proposed technique for RGB color images

The quality factor of the MJPEG is varied and produce multiple quality videos from the same video. Then compression ratio, MSE and PSNR of those MGPEGs are computed. In Figure 6.9, it is seen that for a fixed compression ratio, proposed technique provides better video quality.

For example, with 50% compression, PSNR 51 and 43.5 dB are seen for proposed approach and MGPEG respectively. For the same compression (50%), MSE 0.3 for proposed scheme and 2.9 for MJPEG are found. On the other hand, for a target PSNR of 42 dB, the proposed scheme can compress 77% which is much higher than MJPEG (58%).

6.4.4 Bandwidth savings

In Figure 6.10, estimated bandwidth savings for the different threshold is shown. It is seen that higher threshold introduces noise for reconstructed videos. But increases threshold after certain level cannot improve the bandwidth savings. With this finding, this case study concludes that threshold value 4 or 5 is a good choice in terms of energy consumption and quality of the video.

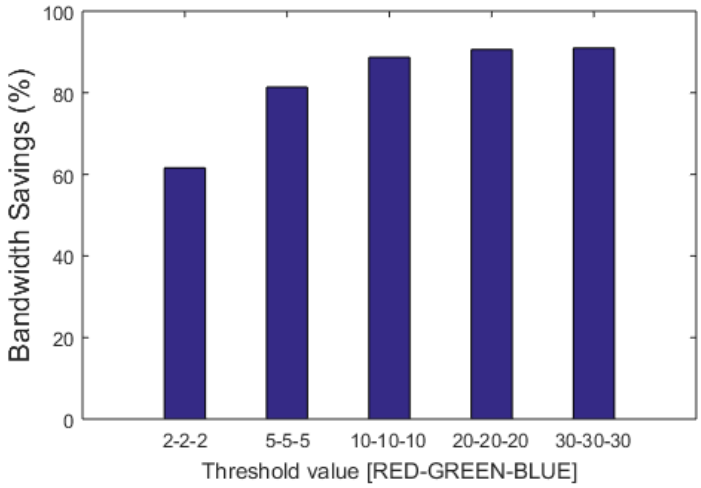


Figure 6.11: Performance comparison of required bandwidth for different value of threshold

6.5 Advantages Towards IoT and IoVT

The proposed technique contributes IoT and IoVT in several reasons. The experiments show that the bandwidth consumption can be reduced by implementing this technique. The quality of the images depends on the threshold value. If the threshold value is lower, this technique can produce high quality of images whereas it has an effect on bandwidth consumption in a reverse way. If the threshold value is lower, that would increase the size of the bitstream resulting high bandwidth consumption. But from the experiment, the result shows that threshold value of 5 would

be a good candidate for this technique. The advantages of this technique towards IoVT are as discussed in this chapter.

1) Network Bandwidth - The proposed system can save significant bandwidth. It can save around 80% of bandwidth when threshold value is 5. If the cameras need to send images in 24/7, this bandwidth efficient technique can serve a lot for IoT based video surveillance.

2) Low Latency – As this technique can save a significant amount of bandwidth, from fog to cloud, the time requires for taking the analytical decision from cloud side reduces. Crimes such as breaking doors or locks of a highly secured area need to be detected within a second or some milliseconds. Automatic decision making (such as shutdown the lift) in a critical situation (such as stampede, fire) need to be executed immediately. For this type of operation, the proposed solution gives a tremendous influence in terms of latency by using in Bayer images in surveillance nodes.

3) Real-time interactions - This technique will give a better result for hard real-time operations. As compression is done in fog node, the time gap is very limited for the cloud to making the decision.

6.6 Summary

In summary, in this chapter, a bandwidth efficient color reproduction technique by variable channel skipping that is targeted for IoT based video surveillance application is presented. The proposed IoT based design achieves low-energy consumption as well as can penetrate the issues faced for IoT based video surveillance. Therefore, the proposed system can be a lightweight preference for Fog to the application of video surveillance.

Chapter 7 - Conclusion and Future Work

7.1 Overview

This thesis deals with a topic which falls under a subfield of IoT, namely IoVT. The main idea of this study is to design a conceptual model of video surveillance system under IoVT environment. One of the goals of this thesis is to identify the features and requirements of the next generation video surveillance system. Considering those requirements, IoT based video surveillance system is modeled and the research shows the advantages of this design that help to meet the requirements. The challenges and issues faced by introducing IoT based video surveillance system are discussed and to eliminate those issues, fog computing is introduced and shows how fog can be a good idea to overcome those issues. This proposed model can be an effective solution in the real-time secured video surveillance system to overcome challenges in emerging IoT environment.

The thesis starts with an overview of video surveillance and internet of things. In chapter 2, the evolution of video surveillance system towards intelligent video surveillance (IVS) and the application sector of it in the modern world has been discussed. The Internet of Things and fog computing, their sector of application including cloudIoT in literature are also been reviewed.

In chapter 3, design requirements for IoT-based video surveillance system are presented. The requirements include the basic requirements which most today's surveillance system already have as well as upgraded requirements which rely on video analytics, computer vision technology and recent advancement of machine learning, artificial intelligence and data mining.

In chapter 4, the research includes the establishment of a conceptual framework for IoT based video surveillance system. A bundle of challenges to integrate IoT in video surveillance system are shown. The fog computing concept is adopted in video surveillance system since it is suitable and introduce to eliminate the challenges of IoT-based system.

In chapter 5, a color reproduction technique for IoT-based video surveillance application is presented as a case study which is energy and bandwidth efficient. The most important advantage of the proposed scheme is lower computational and transmission cost for video surveillance node in IoVT environment. As there is no operation introduced other than the converting color space

and transmitting a single component frames from the encoder, the approach is of low complexity. The proposed IoT-based design achieves consumption of low-energy as well as good quality reconstructed video. Sensor nodes can perform more operations with energy constraints by introducing this approach. Therefore, the proposed design can be a lightweight application of video surveillance in IoT environment.

In chapter 6, a new compression methodology was proposed by variable length pixel encoding. This experiment was done both with Bayer images and RGB images. The proposed system can be implemented in sensor node by compressing Bayer image compression and in fog node for RGB image compression. Implementing the proposed method in fog node can be bandwidth and energy efficient solution and can impact in latency issues of IoT based system.

The contributions of this thesis are summarized below:

1. Intelligent requirements of next generation video surveillance system have been investigated.

2. The challenges for IoT based video surveillance system have been evaluated.

3. It is feasible to implement a fog based video surveillance system to eliminate the issues of IoT based video surveillance system.

4. Energy and bandwidth efficient frame sampling technique for IoT based video surveillance system has been proposed.

5. Energy efficient and low latency based variable length pixel encoding technique for surveillance applications has been proposed.

7.1 Future Works

According to the literature, this is one of the first attempts towards fog-cloud integrated video surveillance system, hence opens a lot of research opportunities in this field. Some of the crucial broad level future research direction includes –

- Task Distribution - In the proposed architecture, fog and cloud collaboratively provide the intelligent analytic requirements of video surveillance system. The distribution of analytic tasks among cloud and fog can be an interesting future research.
- Archiving Methodology - According to the proposed IoT based video surveillance system, video is archived in both fog and cloud. Fog should archive video temporary

basis and cloud should archive for the long term. How this archiving can be done in fog and cloud in bandwidth efficient manner, can be a good future research.

- Fog device - In current IoT system, fog can be implemented in a network device. The sophisticated video surveillance system may require dedicated fog device. The design and implementation of a dedicated fog device for video surveillance can be another good research area.
- Security Mechanism - From the big picture of IoT applications, it is evident that the new technological opportunities have personal, organizational, cultural and social implications. There can be many threats and attacks faced by video surveillance system. Hence, design and implement a security mechanism against different types of threats and attacks can also be a future research topic.
- Prototype Implementation - In this thesis, a fog-cloud based video surveillance system has been proposed. It shows that this proposed system can be a good solution to overcome IoT based issues. However, the implementation of this system for experiment requires a lot of analytic application and IoT protocols. Implementing the whole system and evaluate the performance and challenges will be also a good future research area.

REFERENCES

[1] V. Gouaillier and A. Fleurant, "Intelligent video surveillance: Promises and challenges," *Technological and commercial intelligence report, CRIM and Technôpole Defence and Security,* vol. 456, p. 468, 2009.

[2] V. Gouaillier, "La vidéosurveillance intelligente: promesses et défis," *Rapport technique, TechnoPole Defense and Security, CRIM,* 2009.

[3] T. Huang, "Surveillance video: The biggest big data," *Computing Now,* vol. 7, no. 2, pp. 82-91, 2014.

[4] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in *Proceedings of the first edition of the MCC workshop on Mobile cloud computing*, 2012, pp. 13-16: ACM.

[5] I. Stojmenovic and S. Wen, "The fog computing paradigm: Scenarios and security issues," in *Computer Science and Information Systems (FedCSIS), 2014 Federated Conference on*, 2014, pp. 1-8: IEEE.

[6] F. Bonomi, R. Milito, P. Natarajan, and J. Zhu, "Fog computing: A platform for internet of things and analytics," in *Big Data and Internet of Things: A Roadmap for Smart Environments*: Springer, 2014, pp. 169-186.

[7] S. Yi, C. Li, and Q. Li, "A survey of fog computing: concepts, applications and issues," in *Proceedings of the 2015 Workshop on Mobile Big Data*, 2015, pp. 37-42: ACM.

[8] I. Stojmenovic, S. Wen, X. Huang, and H. Luan, "An overview of fog computing and its security issues," *Concurrency and Computation: Practice and Experience,* vol. 28, no. 10, pp. 2991-3005, 2016.

[9] S. Leader, "Telecommunications handbook for transportation professionals: The basics of telecommunications," 2004.

[10] J. Hourdakis, T. Morris, P. Michalopoulos, and P. Sinha, "Advanced portable wireless measurement and observation station," 2005.

[11] N. Luo, "A wireless traffic surveillance system using video analytics," University of North Texas, 2011.

[12] C. Hartung, R. Han, C. Seielstad, and S. Holbrook, "FireWxNet: A multi-tiered portable wireless system for monitoring weather conditions in wildland fire environments," in *Proceedings of the 4th international conference on Mobile systems, applications and services*, 2006, pp. 28-41: ACM.

[13] A. Kawamura, Y. Yoshimitsu, K. Kajitani, T. Naito, K. Fujimura, and S. Kamijo, "Smart camera network system for use in railway stations," in *Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on*, 2011, pp. 85-90: IEEE.

[14] N. Li, B. Yan, G. Chen, P. Govindaswamy, and J. Wang, "Design and implementation of a sensor-based wireless camera system for continuous monitoring in assistive environments," *Personal and Ubiquitous Computing*, vol. 14, no. 6, pp. 499-510, 2010.

[15] F. Nilsson, *Intelligent network video: Understanding modern video surveillance systems*. CRC Press, 2016.

[16] J. Honovich, "Security manager's guide to video surveillance," *V3. IPVideoMarket. info.[online, ebook] Available: http://ipvideomarket. info/book*, 2009.

[17] D. Elliott, "Intelligent video solution: A definition," *Security,* vol. 47, no. 6, 2010.

[18] H. Liu, S. Chen, and N. Kubota, "Intelligent video systems and analytics: A survey," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 3, pp. 1222-1233, 2013.

[19] Cisco. *Internet of things*. Accessed: Oct. 22, 2017. [Online]. Available: https://www.cisco.com/c/en/us/solutions/internet-of-things/overview.html

[20] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer networks*, vol. 54, no. 15, pp. 2787-2805, 2010.

[21] L. Yan, Y. Zhang, L. T. Yang, and H. Ning, *The Internet of things: from RFID to the next-generation pervasive networked systems*. CRC Press, 2008.

[22] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future generation computer systems,* vol.

29, no. 7, pp. 1645-1660, 2013.

[23] L. Da Xu, W. He, and S. Li, "Internet of things in industries: A survey," *IEEE Transactions on industrial informatics,* vol. 10, no. 4, pp. 2233-2243, 2014.

[24] Z. Pang, Q. Chen, J. Tian, L. Zheng, and E. Dubrova, "Ecosystem analysis in the design of open platform-based in-home healthcare terminals towards the internet-of-things," in *Advanced Communication Technology (ICACT), 2013 15th International Conference on*, 2013, pp. 529-534: IEEE.

[25] M. C. Domingo, "An overview of the Internet of Things for people with disabilities," *Journal of Network and Computer Applications,* vol. 35, no. 2, pp. 584-596, 2012.

[26] H. Alemdar and C. Ersoy, "Wireless sensor networks for healthcare: A survey," *Computer Networks,* vol. 54, no. 15, pp. 2688-2710, 2010.

[27] I. Plaza, L. MartíN, S. Martin, and C. Medrano, "Mobile applications in an aging society: Status and trends," *Journal of Systems and Software,* vol. 84, no. 11, pp. 1977-1988, 2011.

[28] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," *IEEE Internet of Things journal,* vol. 1, no. 1, pp. 22-32, 2014.

[29] T. Nuortio, J. Kytöjoki, H. Niska, and O. Bräysy, "Improved route planning and scheduling of waste collection and transport," *Expert systems with applications,* vol. 30, no. 2, pp. 223-232, 2006.

[30] A. Al-Ali, I. Zualkernan, and F. Aloul, "A mobile GPRS-sensors array for air pollution monitoring," *IEEE Sensors Journal,* vol. 10, no. 10, pp. 1666-1671, 2010.

[31] N. Maisonneuve, M. Stevens, M. E. Niessen, P. Hanappe, and L. Steels, "Citizen noise pollution monitoring," in *Proceedings of the 10th Annual International Conference on Digital Government Research: Social Networks: Making Connections between Citizens, Data and Government*, 2009, pp. 96-103: Digital Government Society of North America.

[32] X. Li, W. Shu, M. Li, H.-Y. Huang, P.-E. Luo, and M.-Y. Wu, "Performance evaluation of vehicle-based mobile sensor networks for traffic monitoring," *IEEE transactions on vehicular technology,* vol. 58, no. 4, pp. 1647-1653, 2009.

[33] S. Lee, D. Yoon, and A. Ghosh, "Intelligent parking lot application using wireless sensor networks," in *Collaborative Technologies and Systems, 2008. CTS 2008. International Symposium on*, 2008, pp. 48-57: IEEE.

[34] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer networks*, vol. 54, no. 15, pp. 2787-2805, 2010.

[35] B. Karakostas, "A DNS architecture for the internet of things: A case study in transport logistics," *Procedia Computer Science*, vol. 19, pp. 594-601, 2013.

[36] E. Qin, Y. Long, C. Zhang, and L. Huang, "Cloud computing and the internet of things: Technology innovation in automobile service," in *International Conference on Human Interface and the Management of Information*, 2013, pp. 173-180: Springer.

[37] P. Mell and T. Grance, "The NIST definition of cloud computing," 2011.

[38] A. R. Biswas and R. Giaffreda, "IoT and cloud convergence: Opportunities and challenges," in *Internet of Things (WF-IoT), 2014 IEEE World Forum on*, 2014, pp. 375-376: IEEE.

[39] J. Zhou *et al.*, "Cloudthings: A common architecture for integrating the internet of things with cloud computing," in *Computer Supported Cooperative Work in Design (CSCWD), 2013 IEEE 17th International Conference on*, 2013, pp. 651-657: IEEE.

[40] K.-D. Chang, C.-Y. Chen, J.-L. Chen, and H.-C. Chao, "Internet of things and cloud computing for future internet," *Security-enriched urban computing and smart grid*, pp. 1-10, 2011.

[41] K. Lee, D. Murray, D. Hughes, and W. Joosen, "Extending sensor networks into the cloud using amazon web services," in *Networked Embedded Systems for Enterprise Applications (NESEA), 2010 IEEE International Conference on*, 2010, pp. 1-7: IEEE.

[42] F. Gao, "Vsaas model on dragon-lab," *International Journal of Multimedia & Ubiquitous Engineering*, vol. 8, no. 4, 2013.

[43] A. Prati, R. Vezzani, M. Fornaciari, and R. Cucchiara, "Intelligent video surveillance as a service," in *Intelligent Multimedia Surveillance*: Springer, 2013, pp. 1-16.

[44] Data Center Companies. Accessed: Oct. 22, 2017. [Online]. Available: https://www.datacenters.com/directory/companies

[45] Cisco. *Cisco Delivers Vision of Fog Computing to Accelerate Value from Billions of Connected Devices*. Accessed: Oct. 22, 2017. [Online]. Available: https://newsroom.cisco.com/pressrelease-content?type=webcontent&articleId=1334100

[46] M. Aazam and E.-N. Huh, "Fog computing: The cloud-iot\/ioe middleware paradigm," *IEEE Potentials*, vol. 35, no. 3, pp. 40-44, 2016.

[47] *OpenFog Consortium*. Accessed: Oct. 22, 2017. [Online]. Available: https://www.openfogconsortium.org

[48] *OpenFog Reference Architecture*. Accessed: Oct. 22, 2017. [Online]. Available: https://www.openfogconsortium.org/ra/

[49] Cloudlet Applications. Accessed: Oct. 22, 2017. [Online]. Available: https://www.akamai.com/us/en/products/web-performance/cloudlets/

[50] *Extending Intelligence to the Edge*. Accessed: Oct. 22, 2017. [Online]. Available: https://itpeernetwork.intel.com/extending-intelligence-to-the-edge/

[51] M. Chiang and T. Zhang, "Fog and IoT: An overview of research opportunities," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 854-864, 2016.

[52] S. Sarkar, S. Chatterjee, and S. Misra, "Assessment of the suitability of fog computing in the context of internet of things," *IEEE Transactions on Cloud Computing*, 2015.

[53] M. Mukherjee *et al.*, "Security and Privacy in Fog Computing: Challenges," *IEEE Access*, 2017.

[54] S. Sarkar and S. Misra, "Theoretical modelling of fog computing: a green computing paradigm to support IoT applications," *IET Networks*, vol. 5, no. 2, pp. 23-29, 2016.

[55] B. Varghese, N. Wang, D. S. Nikolopoulos, and R. Buyya, "Feasibility of Fog Computing," *arXiv preprint arXiv:1701.05451*, 2017.

[56] K. Dantu, S. Y. Ko, and L. Ziarek, "RAINA: Reliability and Adaptability in Android for Fog Computing," *IEEE Communications Magazine*, vol. 55, no. 4, pp. 41-45, 2017.

[57] P. Varshney and Y. Simmhan, "Demystifying Fog Computing: Characterizing Architectures, Applications and Abstractions," *arXiv preprint arXiv:1702.06331,* 2017.

[58] C. Perera, Y. Qin, J. C. Estrella, S. Reiff-Marganiec, and A. V. Vasilakos, "Fog computing for sustainable smart cities: a survey," *arXiv preprint arXiv:1703.07079,* 2017.

[59] E. K. Markakis *et al.*, "EXEGESIS: Extreme Edge Resource Harvesting for a Virtualized Fog Environment," *IEEE Communications Magazine*, vol. 55, no. 7, pp. 173-179, 2017.

[60] H. Dubey *et al.*, "Fog Computing in Medical Internet-of-Things: Architecture, Implementation, and Applications," in *Handbook of Large-Scale Distributed Computing in Smart Healthcare*: Springer, 2017, pp. 281-321.

[61] A. M. Rahmani *et al.*, "Exploiting smart e-health gateways at the edge of healthcare internet-of-things: a fog computing approach," *Future Generation Computer Systems,* vol. 78, pp. 641-658, 2018.

[62] F. A. Kraemer, A. E. Braten, N. Tamkittikhun, and D. Palma, "Fog Computing in Healthcare–A Review and Discussion," *IEEE Access*, 2017.

[63] R. Roman, J. Lopez, and M. Mambo, "Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges," *Future Generation Computer Systems,* vol. 78, pp. 680-698, 2018.

[64] R. Hasan, S. K. Mohammed, A. H. Khan, and K. A. Wahid, "A color frame reproduction technique for IoT-based video surveillance application," in *Circuits and Systems (ISCAS), 2017 IEEE International Symposium on*, 2017, pp. 1-4: IEEE.

[65] J. Xu, Y. Andrepoulos, Y. Xiao, and M. van Der Schaar, "Non-stationary resource allocation policies for delay-constrained video streaming: Application to video over Internet-of-Things-enabled networks," *IEEE Journal on Selected Areas in Communications,* vol. 32, no. 4, pp. 782-794, 2014.

[66] R. Pereira and E. G. Pereira, "Video streaming considerations for internet of things," in *Future Internet of Things and Cloud (FiCloud), 2014 International Conference on*, 2014, pp. 48-52: IEEE.

[67] Z. He, W. Cheng, and X. Chen, "Energy minimization of portable video communication devices based on power-rate-distortion optimization," *IEEE transactions on circuits and systems for video technology*, vol. 18, no. 5, pp. 596-608, 2008.

[68] M. Malić, D. Dobrilović, and I. Petrov, "Example of IoT platform usage for wireless video surveillance with support of NoSQL and cloud systems," 2016.

[69] U. Jennehag, S. Forsstrom, and F. V. Fiordigigli, "Low Delay Video Streaming on the Internet of Things Using Raspberry Pi," *Electronics*, vol. 5, no. 3, p. 60, 2016.

[70] N. Chen, Y. Chen, Y. You, H. Ling, P. Liang, and R. Zimmermann, "Dynamic urban surveillance video stream processing using fog computing," in *Multimedia Big Data (BigMM), 2016 IEEE Second International Conference on*, 2016, pp. 105-112: IEEE.

[71] P. Corcoran, "Beyond stream processing—A distributed vision architecture for the Internet of Things," in *Consumer Electronics (ICCE), 2016 IEEE International Conference on*, 2016, pp. 168-169: IEEE.

[72] L. Yatziv and G. Sapiro, "Fast image and video colorization using chrominance blending," *IEEE Transactions on Image Processing*, vol. 15, no. 5, pp. 1120-1129, 2006.

[73] T. Welsh, M. Ashikhmin, and K. Mueller, "Transferring color to greyscale images," in *ACM Transactions on Graphics (TOG)*, 2002, vol. 21, no. 3, pp. 277-280: ACM.

[74] T. Horiuchi and S. Hirano, "Colorization algorithm for grayscale image by propagating seed pixels," in *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, 2003, vol. 1, pp. I-457: IEEE.

[75] A. Levin, D. Lischinski, and Y. Weiss, "Colorization using optimization," in *ACM transactions on graphics (tog)*, 2004, vol. 23, no. 3, pp. 689-694: ACM.

[76] V. Korostyshevskiy. *Grayscale to rgb converter*. MATLAB Central file exchange [Online], 2006.

[77] T. Khan, R. Shrestha, M. S. Imtiaz, and K. A. Wahid, "Colour-reproduction algorithm for transmitting variable video frames and its application to capsule endoscopy," *Healthcare technology letters*, vol. 2, no. 3, pp. 52-57, 2015.

[78] J. H. Ko, B. A. Mudassar, and S. Mukhopadhyay, "An energy-efficient wireless video sensor node for moving object surveillance," *IEEE Transactions on Multi-Scale Computing Systems,* vol. 1, no. 1, pp. 7-18, 2015.

[79] L. G. R. Diaz, P. L. C. Burgos, and D. F. Rodriquez, "Video monitoring and security system," ed: Google Patents, 2002.

[80] H. S. Parekh, D. G. Thakore, and U. K. Jaliya, "A survey on object detection and tracking methods," *International Journal of Innovative Research in Computer and Communication Engineering,* vol. 2, no. 2, pp. 2970-2979, 2014.

[81] K. S. Huang and M. M. Trivedi, "Video arrays for real-time tracking of person, head, and face in an intelligent room," *Machine vision and applications,* vol. 14, no. 2, pp. 103-111, 2003.

[82] F. Chen and C. De Vleeschouwer, "Personalized production of basketball videos from multi-sensored data under limited display resolution," *Computer Vision and Image Understanding,* vol. 114, no. 6, pp. 667-680, 2010.

[83] X. Wang, "Intelligent multi-camera video surveillance: A review," *Pattern recognition letters,* vol. 34, no. 1, pp. 3-19, 2013.

[84] D. Comaniciu, F. Berton, and V. Ramesh, "Adaptive resolution system for distributed surveillance," *Real-Time Imaging,* vol. 8, no. 5, pp. 427-437, 2002.

[85] Lorex. *Pan–tilt–zoom camera*. Accessed: Oct. 22, 2017. [Online]. Available: https://www.lorextechnology.com/ptz-cameras/N-n6lzjh.

[86] *The guide to panoramic photography*. Accessed: Oct. 22, 2017. [Online]. Available: https://www.panoramic-photo-guide.com/panoramic-photography.html

[87] A. Sarhan, M. T. Faheem, and R. O. Mahmoud, "A proposed intelligent denoising technique for spatial video denoising for real-time applications," in *Advancing the Next-Generation of Mobile Computing: Emerging Technologies*: IGI Global, 2012, pp. 145-163.

[88] X. Zhang, S. Hu, D. Chen, and X. Li, "Fast covariance matching with fuzzy genetic algorithm," *IEEE Transactions on Industrial Informatics,* vol. 8, no. 1, pp. 148-157, 2012.

[89] G. Diamantopoulos and M. Spann, "Event detection for intelligent car park video surveillance," *Real-Time Imaging,* vol. 11, no. 3, pp. 233-243, 2005.

[90] E. Fellores. *Chilean salmon farm enhances security with intelligent video edge*. Accessed: Oct. 22, 2017. [Online]. Available: https://www.securitymagazine.com/articles/79254-chilean-salmon-farm-enhances-security-with-intelligent-video-edge-1

[91] R. Chang, L. Guan, and J. Burne, "An automated form of video image analysis applied to classification of movement disorders," *Disability and rehabilitation,* vol. 22, no. 1-2, pp. 97-108, 2000.

[92] K. F. MacDorman, H. Nobuta, S. Koizumi, and H. Ishiguro, "Memory-based attention control for activity recognition at a subway station," *IEEE MultiMedia,* vol. 14, no. 2, 2007.

[93] T. Wang, Z. Zhu, and E. Blasch, "Bio-inspired adaptive hyperspectral imaging for real-time target tracking," *IEEE Sensors Journal,* vol. 10, no. 3, pp. 647-654, 2010.

[94] C. Quek, W. Irawan, and E. Ng, "A novel brain-inspired neural cognitive approach to SARS thermal image analysis," *Expert Systems with Applications,* vol. 37, no. 4, pp. 3040-3054, 2010.

[95] C. Tran and M. M. Trivedi, "3-D posture and gesture recognition for interactivity in smart spaces," *IEEE Transactions on Industrial Informatics,* vol. 8, no. 1, pp. 178-187, 2012.

[96] H. Rastgar, M. Ahmadi, and M. Sid-Ahmed, "A stereo vision-based bin picking system using Hopfield neural networks," *Journal of Circuits, Systems, and Computers,* vol. 18, no. 03, pp. 443-463, 2009.

[97] W. Lee, "3D machine vision in IoT for factory and building automation," in *Circuits and Systems (ISCAS), 2017 IEEE International Symposium on*, 2017, pp. 1-1: IEEE.

[98] *Panasonic intelligent video technology whitepaper. INTELLIGENT VIDEO ANALYTICS LEVERAGE REAL-TIME VIDEO SURVIELLANCE DATA FOR POWERFUL BUSINESS DECISIONS*. Accessed: Oct. 22, 2017. [Online]. Available: https://www.sdmmag.com/ext/resources/files/White_Papers/Panasonic_Intelligent-Video-Technology_Whitepaper.pdf

[99] E. Soyak, S. A. Tsaftaris, and A. K. Katsaggelos, "Low-complexity tracking-aware H. 264 video compression for transportation surveillance," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 21, no. 10, pp. 1378-1389, 2011.

[100] T. Khan, and K. A. Wahid, "Low-complexity colour-space for capsule endoscopy image compression," *Electronics letters,* 47.22 (2011): 1217-1218.

[101] *Video Surveillance Online Repository*. Accessed: Oct. 22, 2017. [Online]. Available: http://www.openvisor.org/video_videosInCategory.asp?idcategory=20