Enabling Green Video Streaming over Internet of Things

(1^{st} Quarter Deliverable)

Dr. Ghalib Asadullah Shah (PI)

Next-generation Wireless Networking (NWN) Lab,
Al-Khawarizmi Institute of Computer Sciences,
University of Engineering and Technology, Lahore

04-03-2014

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction to Internet of Things

The explosive and ubiquitous adoption of Wireless Sensor Networks (WSNs) in applications such as environmental monitoring, agriculture, healthcare, and smart buildings, is due to its low cost sensor technology. WSN have enabled by large-scale networks of small devices capable of harvesting information from the physical environment, performing simple processing on the extracted data and transmitting it to remote locations [6]. Generally, WSNs are based on heterogeneous proprietary and non-proprietary solutions. To deploy large virtual WSNs with heterogeneous systems and enable interoperability between them requires hardware as well as application compliant solutions.

Proprietary or closed system based WSNs can be realized as connectivity islands, and they require some application definite gateways to export data through Internet to the destination nodes. However, to enable communication between heterogeneous technologies based WSNs, avoiding complex application-specific conversions at gateways; recently the Internet Protocol (IP) is promoted to connect WSNs devices to the Internet. Thereby, tiny network devices i.e. sensors, actuators etc are enabled with unique IP addresses and network interfaces making them smart objects based on open standards to build an 'Internet of Things' (IoT).

The concept of IoT carries enormous number of ideas that is why overlapping definitions are given for IoT. As identified by Atzori et. al. [7], IoT can be realized in three paradigms - internet-oriented (middleware), things oriented (sensors) and semantic-oriented (knowledge). Reason being the interdisciplinary characteristics of the IoT the demineralization is helpful, however the true potential of IoT can only be realized when these three paradigms are combined. The most widely accepted definition of IoT is given by Radio-frequency identification (RFID) group as; the worldwide network of interconnected objects uniquely addressable based on standard communication protocols.

On the other hand, Cluster of European research projects on the IoT [8] state

that, Things are active participants in business, information and social processes where they are enabled to interact and communicate among themselves and with the environment by exchanging data and information sensed about the environment, while reacting autonomously to the real/physical world events and influencing it by running processes that trigger actions and create services with or without direct human intervention. Forrester defines IoT as, a smart environment uses information and communications technologies to make the critical infrastructure components and services of a city administration, education, healthcare, public safety, real estate, transportation and utilities more aware, interactive and efficient [9].

Enabling simple embedded WSNs devices communicating to other devices connected through Internet infrastructure has promoted more conceptual designs, visions and applications of the IoT are depicted in [10]-[11]. There are enormous applications that are possible because of IoT. For individual users, IoT brings useful applications like home automation, security, automated devices monitoring and management of daily tasks. For professionals, automated applications provide useful contextual information all the time to help on their works and decision making. For Industrialists Internet enabled sensors and actuators operations can be rapid, efficient and more economic. Managers who need to keep eye on many things can automate tasks connection digital and physical objects together.

1.2 Requirements and Challenges for IoT

The IoT enables communication between heterogeneous system devices through Internet, which comes with big challenges such as dealing with large amounts of information, queries, and computation. For this reason, IoT requires communication standards as Hyper Text Transfer Protocol (HTTP) [12], Internet Protocol (IP) [13], Transmission Control Protocol (TCP) [14], to run operations such as information data processing, streaming, filtering, aggregation and data mining. However, direct implication of these protocols for IoT is not straight forward, since these protocols are not designed considering energy constrained devices. On the contrary, IoT devices have strict energy limitation to enable tiny devices to get connected to the Internet; these objects are usually powered by batteries or through energy-harvesting. The existing protocols i.e. HTTP, IP, TCP, consume significant amount of energy due to the redundant data transmissions, protocol overheads, headers transmission, acknowledgment packets for higher layers to ensure reliability etc. These communication protocols are not optimized for low power communication as required in IoT, therefore adaptation of these mechanisms is infeasible in their current form and structure.

Unlike traditional networks with device dependent systems like WSNs, RFID [15], Machine-to-Machine (M2M) [16], [17], IoT enables Internet-connected heterogeneous devices to communicate with one another. To achieve a practical realization of IoT, following three requirements must be met by the communication

mechanism or communication stack employed for IoT:

1.2.1 Energy Efficiency

Most of the devices in a practical IoT have a limited available power, from energy sources like batteries or energy harvesting. Since, replacement or recharging of these batteries is in practical. Therefore, communication mechanism should ensure low power consumption in its operation.

1.2.2 Highly Reliable

Current communication stack for Internet uses error detection, retransmissions and flow control etc to provide best effort transport quality of service. These functionality are applied on different layers to ensure reliability; however, the performance of the overall communication stack is poor in terms of efficiency i.e. associated overheads, energy wastage, etc. For seamless deployment of IoT same level of reliability as enjoyed by the Internet must be ensured, however at the maximum achievable efficiency.

1.2.3 Internet-Enabled

The Internet architecture exhibits the ability to facilitate bidirectional communication between end to end devices. Similarly, it is of vital importance in case of IoT that not only the end devices communicate or share information to other entities but also it should be enabled the other way around. In addition, Internet enables devices with different lower layer technologies to communicate through a common language, IP. As IoT is expected to connect nearly all the 'Things' with all other 'Things', therefore enabling the devices with IP technology is a must for IoT, so that the 'Things' can be identified and be communicated by using the uniquely identifiable IP addresses [18]. As a consequence of this, standardization for new communication protocols is required for meeting the specific needs of IoT.

Various efforts have been made by IETF and IEEE to address these requirements in order to realize Internet of things. Chapter **??** discusses the recommended protocols for IoT in details.

1.3 Internet of Multimedia

Recently due to the availability of low cost multimedia devices i.e. CMOS cameras and microphones, the Wireless Multimedia Sensor Networks (WMSNs) have gain lots of attraction. In WSMNs ubiquitously distributed low-power low cost devices communication to retrieve multimedia content from the physical environment in the form of video and audio streams, still images, or scalar sensor data. Conventionally, these tiny devices report data to a centralized network manager

or gateway or application server device, which is responsible to store and fuse the data coming from different devices. Later, based on the information in this data various processes can be triggered by the coordinator device locally or it can be reported to some other entity possibly through Internet.

The Internet of Multimedia (IoM) is an enhancement to the IoT, whose prime objective is to enable video streaming as part of the realization of IoT. In IoM, resource constraint low-power low-cost heterogeneous multimedia devices can be connected and each device can be globally accessible by a unique IP address with the same spirit as of the computers and other networking devices connected through the Internet. This approach enables a wide range of applications in the areas of home and building automation, factory monitoring, smart cities, transportation, smart grid and energy management [19]. For example,

- Multimedia surveillance sensor networks to deal with crime and terrorist attacks
- Traffic avoidance, enforcement and control systems to offer traffic routing advice, etc
- Advanced health care delivery enabling medical centers to remotely monitor patients
- Person locator service to locate missing persons, or identify criminals and terrorists
- Industrial process control to automatically inspect any defect in products, etc

1.4 Requirements and Challenges for IoM

IoM implication in is true color requires resource constraints multimedia devices to be able to communicate through the Internet with various heterogeneous intermediate devices. This requirement comes with same challenges as in case of IoT such as dealing with large amounts of information, queries, and computation. The delivery mechanism of the current Internet architecture, offers 'Best-Effort' transmission with no guarantee of successful packet delivery or the time of its delivery [20]. Despite the fact that 'Best-Effort' transmission mechanism is suitable for non-real time data communication which is delay tolerable nut required error-free delivery, whereas this approach is not appropriate for time-constraint multimedia data such as video and audio.

The multimedia content being communicates over the network requires different traffic requirements as compared to data communication such as bandwidth, delay, jitter and reliability. These network performance requirements are referred as the Quality of Service (QoS), which represents the level of user experience needed. For example, how fast is the data transmission, how much is the delay at

the receiver, what is the probability of correct data reception at the receiver, what is the probability of the transmitted data to be lost, etc.

The real-time multimedia is continuous in nature which may or may not be delay tolerant depending upon the audio/video application i.e. streaming or interacting. For example, a video camera device generates 25 to 30 frames per second from the optical data. These frames are transmitted in the form of packets require specific bandwidth capacity directly related to video compression being done at the transmitting device. Compressing video to enable a certain level of quality over a low capacity channel offering low data rate, increases encoder complexity and drains energy. Some popular video compression techniques [21] are given in Table 4 below. On one side energy constraint IoM device exhibit limited bandwidth capacity, yet enabling a good video quality requires high compression which is also infeasible due to high energy consumption. Thus, there is a tradeoff between the achievable compression and the user experience for a specific energy constraint. In addition, resource constraint multimedia devices require implementation specific QoS, which is realizable with network metrics i.e. delay, latency, jitter, etc. Thus, IoM requires communication standards such as HTTP [12], IP [13], TCP [14], etc. However, direct implication of these protocols for IoM is not straight forward, since these protocols are not designed considering energy constrained devices.

On the contrary, IoM multimedia devices are realized as low-cost, battery powered, low-complexity devices and the existing protocols i.e. HTTP, IP, TCP, consume significant amount of energy due the redundant data transmissions, protocol overheads, headers transmission, acknowledgment packets for higher layers to ensure reliability etc. Since, these communication protocols are not optimized for low power communication as required in IoM, therefore adaptation of these mechanisms is infeasible in their current form and structure.

The current standardization activities of providing Internet-connectivity to 'Things' [**?**] are not focused to address the challenges of provisioning multimedia objects over Internet of Things. The main obstacles of realizing IoM that enables Internet access to WMSN are limited available power, limited available capacity, and heterogeneity of multimedia devices. Many researchers have investigated a variety of techniques to limit the power consumption of wireless multimedia sensor networks. However, these issues have not been addressed considering WMSNs based on IoM architecture.

1.4.1 Energy Efficiency

In practical WMSNs the multimedia devices are small sized objects and occupy a limited amount of power, which they have to utilize efficiently to increase network life time. Multimedia transmission consumes an excessive amount of energy that needs to be dealt carefully, since the existing power-saving solutions developed for WMSNs are not tailored to the multimedia support required for IoM architecture. In addition to the energy efficiency, multimedia streaming needs to incorporate

green communication solution for sake of limiting the carbon footprints in the environment, particularly when the multimedia devices are installed at large scale as envisioned in IoT.

1.4.2 High Data Rate

The current standardization activities of providing Internet access to things are not focused to address the challenges of provisioning multimedia objects over IoTs. Things in IoT are commonly assumed to be small objects generating periodical or event driven scalar data in the order of few tens of octets. Although ZigBee or IEEE 802.15.4 standard is devised for low power devices to operate at short range for IoT, it is not possible to employ it for wireless multimedia sensors, present in IoM, due to its low rate in the range of 50-250 kbps. However, the most widely video encoding standard used for small embedded devices is H.264 that produces bit-rate even at low resolution (384 kbps at 320x240 at 20 frames per second) much higher than the maximum data rate of 250kbps in IEEE 802.15.4. Moreover, the video frame size is in the order of KB that would require a large number of physical layer frame transmissions and eventually incurring much higher overhead that cannot be efficiently used for multimedia delivery. Thus, it will prevent the use of multimedia devices to connect through IoT defined protocols and specification.

1.4.3 Internet Connectivity

IoT aims to enable every device or thing to be accessed or connected to Internet for a variety of reasons. The past standardization efforts for low-power wireless networking (e.g., ZigBee [**?**], Z-Wave [22], and WirelessHART are based on ad hoc solutions and believed that IP was too large and ill-suited for the needs of WSNs. Without a common network layer, these systems require an application layer gateway to communicate with other systems or even the wider Internet. The IP for Smart Objects (IPSO) Alliance promote the Internet Protocol as the network technology for connecting Smart Objects around the world with the help of IPv6 [18]. Through IP adaptation and incorporation by low power network standards like IEEE 802.15.4 will enable deployment of IoT paradigm [18], but the latency constraint of multimedia devices have been overlooked in this effort. Thus, enabling small multimedia sensor devices to efficiently communicate with heterogeneous devices is still an open issue.

1.4.4 Adaptive Video Encoding

Considering the heterogeneity of devices from low power CMOS camera in WMSN to high resolution modern digital camera, the video encoder must be adaptive to the hardware and network resources of the device to process and stream in real-time with acceptable video fidelity. The current video encoders are focusing more on

the reduction of video coding rate but ignore the device processing capability that hinders its implementation on the low power multimedia devices. Thus developing such an encoder that is aware of the devices resources and network bandwidth is the pressing need of IoM.

1.5 Wireless Connectivity for IoM

Multimedia applications require high data rates. For real-time multimedia communication, in particular, data transmission has to be maintained at high rates. IEEE 802.11 standard provides desired data rates and therefore is ubiquitously suggested in the literature as the standard for wireless connectivity. Wi-fi offers plethora of benefits. Apart from the recently developed power-efficient Wi-Fi models offering battery lifetime spanning over multiple years[23], reuse of existing Wi-Fi infrastructure offers cost savings and faster deployments. Wi-Fi devices have the advantage of native IP-network compatibility, Wi-Fi also provides differentiated quality of service. Similarly, to provide reliable communication Wi-Fi Protected Access (WPA/WPA2) provides link layer encryption and authentication support. the random access mechanism and higher data rate support provides support to improve latency, end-to-end delay, and jitter etc, in other words provides better user experience for multimedia communication. In addition, IEEE 802.11 provides reliable communication using acknowledgments and transmitted packets ordering support. Moreover, it can support transmission range corresponding to the view range of video devices that makes it a practical choice. Furthermore economy of scale is another important advantage of Wi-Fi with an expected 22 percent annual growth rate between 2010 and 2015 [24]. For these reasons, Wi-Fi has already been widely accepted for many commercial off-the-shelf video devices, which are largely deployed for video surveillance and monitoring applications, making it a good candidate for IoM.

1.6 Video Encoding

The basic objective of all video coding encoding algorithms is highly specified to achieve a comprised balance with complexity of the methodical solution and optimizing the resultant efficiency that includes transporting the image either viva single route or through distributed resources. Typically video codec are video coding standards to achieve highest video quality in lesser time. A video stream is a frame of picture split from video format. The high-computational video complexity and real-time requirements of video systems represent the main challenges to overcome on the development of efficient encoder solutions. Generally it is considered important if the encoding phase complexity scalability is proportional to temporal and spatial resolution.

Compression scheme	Comments
MPEG-I	Used to produce VCR NTSC (352 x 240) quality video compression to be stored on CD-ROM (CD-I and CD-Video format) using a data rate of 1.2 Mbps. Uses heavy down-sampling of images as well as limits image rate to 24-30 Hz to achieve this goal.
MPEG-II	More generic standard for a variety of audio-visual coding applications and supports error-resilience for broadcasting. Supports broadcast-quality video compression (DVB) and High Definition Television (HDTV). MPEG-2 supports four resolution levels: low (352 x 240), main (720 x 480), high-1440 (1440 x 1152), and high (1920 x 1080). The MPEG-2 compressed video data rates are in the range of 3-100 Mbps.
MPEG-IV	Supports low bandwidth video compression at data rate of 64 Kbps that can be transmitted over a single N-ISDN B channel. MPEG-4 is a genuine multimedia compression standard that supports audio and video as well as synthetic and animated images, text, graphics, texture, and speech synthesis.
H.261	Supports video communications over ISDN at data rates of px64 Kbps. It relies on intra and inter-frame coding where integer-pixel accuracy motion estimation is required for inter mode coding
H.263	The H.263 standard is aimed at video communications over POTS and wireless networks at very low data rates (as low as 18-64 Kbps). Improvements in this standard are due to the incorporation of several features such as half-pixel motion estimation, overlapping and variable blocks sizes, bi-directional temporal prediction, and improved variable-length coding options.

Figure 1.1: Video Compression Techniques.

There have been many advancement in the development of video coding standard, major achievement was the well known standard H.262/MPEG-2. The standard was the outcome of the partnership between Moving Pictures Experts Group (MPEG) and Video Coding Experts Group (VCEG) [25]. VCEG is former and attentive towards conventional video coding goals such as packet-loss and compression where as MPEG is ambitious towards synthetic-natural hybrid coding and object oriented video coding. Further one from there, the evolution picked up and resultantly published new standard H.264/MPEG-AVC which gave 50% efficiency compared with previous standards. The design of H.264 have Fidelity Range Extensions (FRExt), Multiview Video Coding (MVC) extension and Scalable Video Coding (SVC) extension. After the introduction of H.264, in 2010 H.265 was published from the merging of two video coding experts namely ISO MPEG and ITU VCEG, establishing a Joint Collaborative Team (JVT-VC). H.265 is a compatible video coding standard, that manages to be able to work for every h.24x series applications. Besides JVT-VC, other companies have also introduced their own particular video coding standards, few of them are listed below.

HEVC/H.265
High Efficiency Video Coding is advanced video coding standard from previous standard H.264/AVC [26]. In 2003 AVC was published and used in HD television. H.265 increases compression efficiency by two times and also due to its compatibility with every other video standard application it would be replacing H.264. Later in the report H.265 structure have been discussed in detail.

VP9
VP9 is an open source developed video coding standard, which offers saving in network usage, GOOGLE [27] have developed this standard considering their networking issues specifically regarding online video streaming. This standard also have adopted some of similar characteristics of previous video coding standard "AVC". The major difference is instead of macroblock it uses *"super block"* which easily could split adaptively. This technique improves motion vector prediction.

Daala
Daala[28] is developed independently from characteristics shared by H.265 and VP9. It is highly based on hybrid coding, which might overtake the video coding industry. It is patent free and considers lapped transformation rather than block-based DFT filtering and intra prediction in frequency domain.

1.7 Energy Efficient Video Acquisition

Video Encoders offer compression ratios of hundreds of times. This causes a tremendous reduction in transmission rate with tolerable impact on quality. Video is encoded once at the transmitter and decoded each time it is played back at the receiver. Thus, all the existing video compression techniques are designed to have very complex encoders and simple decoders. The complexity of the encoders make them energy hungry. Since the video nodes in IoM have severe battery limitations, IoMs require encoders to be as simple as possible while shifting all the complexity to the decoders. Techniques and methods from a recent and vibrant area of compressive sensing are leveraged in this regard. These methods enables acquisition of video frames which are already compressed as opposed to compressing them after acquisition. The aim therefore is is to reduce video encoder complexity(both space and time), by the use of compressed sensing[6, 8], while maintaining fairly low bit-rates (for transmission). In wireless multimedia sensor networks video quality largely depends on the available band width[7]. So we aim to gain control over the amount of raw data sensed and dynamically relate the quantity to the available battery power, and the available bandwidth (in the dynamic network environment). The key idea is to make a hybrid codex, based on the theory of compressed sensing, while maintaining some useful features of existing codex[29, 30], such as H.264.

The compressive sensing (CS) paradigm provides an efficient image acquisition technique through simultaneous sensing and compression. In CS, a number of random projections of the image are sensed as the compressed version of the image, thus leading to a faster image acquisition system[6]. Since the imaging philosophy in CS imagers is different from conventional imaging systems, one need to find a way to read out not the pixel values themselves but a set of random linear combinations of pixel values. This random linear sampling has been im-

plemented either in the optical domain or on-chip in the circuit domain[8]. In the subsequent sections we will look through a few of the implementations, their advantages/disadvantages, particularly so in relation to the back to back implementation of CS and H.264. The aim of this architecture is to significantly reduce the encoder complexity (both in terms of time and space complexities), while not compromising on the video quality.

Modern video codecs such as H.264, use macro blocks (16x16, 4x4, e.tc) to compress the video[31]. Now if CS is applied on the whole pixel array the resulting CS data will no more be segregable into macro blocks, with any useful spatial information. Consequently H.264 will not be applicable. On the other hand if, for instance, CS is applied on 16x16 macro blocks of the pixel array, and the compression ratio is kept at 1/16 (i.e. 16 measurements per macro block), 4x4 CS macro blocks are obtained. Some modified (simplified) form of H.264 can be applied on these CS macro blocks for further compression, before transmission.

1.8 Roadmap

This report comprises two parts. After the introduction in chapter 1, chapters 2 covers the existing stack of protocols proposed for IoT in detail. Next, the overview of IEEE 802.11 for IoM is provided in Chapter 3 to give an insight for the suitability of WI-Fi in IoM and its exploitable parameters for energy efficiency in low power multimedia devices. Chapter 4 lays down different power saving mechanisms to enhance the existing IEEE 802.11 MAC with new network configuration protocol to adapt in IoM. In the second part, an overview of existing commonly used video encoding algorithm is provided in Chapter 5 to explore its implications in low power devices. This is followed by the new proposed compressive sensing based video encoding algorithm in Chapter 6. Finally, Chapter 7 describes the hardware architecture for implementation of the green camera node and the demonstration of the project.

Chapter 2

IoT Protocol Stack

IEEE and IETF have been working since 2003 to design appropriate communication protocols for next generation networks. However, for IoT the most widely accepted standard is IEEE 802.15.4 [32], which defines both Physical Layer (PHY) and Medium Access Control (MAC) for low power communication that is well suited for IoT. The PHY and MAC defined by IEEE 802.15.4 formed the basis of ZigBee 1.0 and later for ZigBee 2006 [33]. The single-channel characteristic of IEEE 802.15.4 MAC is unreliable in multi-hop network scenarios. Dust Networks [34] introduced a channel hopping protocol named Time Synchronized Mesh Protocol (TSMP) [35] to mitigate multi-path fading which was later adopted by WirelessHART [36]. This TSMP inspired IEEE 802.15.4e working group to design a Time Synchronized Channel Hopping (TSCH) protocol, which later integrated into the IEEE 802.15.4 in 2010. To facilitate low power based devices to get connected to the Internet multiple IETF working groups, proposed low power based communication protocols i.e. 6LoWPAN [37] as a convergence layer, ROLL RPL [38] as a routing protocol, and CoAP [39] as an application layer protocol. These protocols have been considered as key technologies in realization of IoT; in the later sections we will discuss the functionality of each of these protocols.

2.1 Physical Layer

In this section we discuss the importance of a low power based PHY and its requirement to enable energy efficient communication between smart devices in the IoT.

2.1.1 Low-Power Radio Hardware

A PHY receives data from higher layers in binary form; this data is converted into an analog signal and encoded using a specific modulation scheme. This signal is then amplified up to a certain power level and transmitted over the air. The

transmission power of a typical low-power radio is 0 dBm or 1 mW. The wireless channel environment causes loss in the signal energy due to the path loss, fading, shadowing, interference etc. The lowest power level above which a radio can receive the transmitted signal is referred as receiver sensitivity, i.e. a ?90 dBm receiver sensitivity radio can effectively demodulate a signal as weak as 1 pW.

The transmission or reception of signals involve modulation, demodulation, signal amplification etc, these processes consume significant amount of energy, making the radio the major power draining When on, the modulator, demodulator, and PA, all draw substantial amounts of current, making the radio the most power-hungry part of a device. However, when inactive radio consumes no energy, therefore a communication protocol must reduce the radio on/active time or duty cycle to as low as possible (typically lower than 1%) while assuring reliable data communication. The transmission power of typical radios is between -50 dBm to +5 dBm, 0 dBm being the default value. A radio draws same amount of current in listening (idle state) or while receiving a transmitted signal. On the other hand, a radio draw maximum amount of current during the transmission of the signal. Therefore, for energy efficient operation the duty cycle of the radio must be decreased. A list of data sheet numbers for commercially available low-power radios from different vendors is given in Table 2.1.

Generally, the wireless sensors are battery-powered, and replacing these batteries is impractical in most scenarios. To increase the life time of the node or in turn the life time of the network, an energy efficient protocols are required which either adopt a lower duty cycle or utilized a very low-power consuming radio to make the whole operation energy efficient. Let assume a sensor node is powered by a pair of AA batteries, holding 3000 mAh of charge. Then in this case if sensors node uses the AT86RF231 (which draws ? 13 mA when active) and a protocol maintaining a 100% duty cycle then the batteries will be exhausted in 3000 mAh/13 mA = 230 h, or 10 days. However, for the same device, if a protocol with a 1% duty cycle is used then the lifetime will be 32 months. In addition, replacing AT86RF231 with the lower-power LTC5800 (which draws ? 5 mA when active), increases the lifetime to 3000 mAh/(5 mA1%) = 60000 h, or 7 years.

2.1.2 IEEE 802.15.4

IEEE 802.15.4 [32] is one of the most well-known and widely accepted standards for low-power radio technology. The first version of the standard came in 2003, which defined both PHY and MAC layer specifications. Later in 2006, a revised version of the standard came. Many IEFT working groups i.e. IEEE 802.15.4e, are currently working on improving the standard. The IEEE 802.15.4 defined PHY specifications give a well suited trade-off between energy-efficiency, transmission range, and data rate, which makes it most feasible in small to medium sized Wireless Personal Area Networks (WPAN). IEEE 802.15.4 standard defines multiple PHY layers, among which the most commonly employed operating band that is the 2.4?2.485 GHz

Vendor	Product	Sensitivity [dBm]	Transmit current [mA] @ 0 dBm	Receive current [mA]
Atmel	AT86RF231 [a]	-101	14.0	12.3
Dust Networks/ Linear Tech.	LTC5800 [b]	-91	5.4	4.5
Ember	EM357 [b]	-100	27.5	25.0
Freescale	MC13233 [b]	-94	26.6	34.2
Microchip	MRF24J40 [a]	-95	23.0	19.0
NXP/Jennic	JN5148 [b]	-95	15.0 (1.8 dBm)	17.5
Texas Instr.	CC2520 [a]	-98	25.8	18.8

Figure 2.1: Power Consumption of IEEE 802.15.4 Compatible Devices.

frequency band, which is an unlicensed band.

IEEE 802.15.4 PHY layer uses Offset-Quadrature Phase-Shift Keying (O-QPSK) modulation with a support of 2 Mbps PHY data rate. However, to achieve additional robustness Direct Sequence Spread Spectrum (DSSS) is used which encodes every 4 bits as 32 bits, thus reducing the data rate to 250 kbps. IEEE 802.15.4 also defines 16 frequency channels, located every 5 MHz between 2.405 GHz and 2.480 GHz. The channels themselves are only 2 MHz wide, thus they are non-interfering orthogonal channels. The radio can subjectively send and receive on any of those channels, and an IEEE 802.15.4 based radio is able to switch channels in less than 192 μsecs.

When a radio sends a packet, it starts by transmitting a physical preamble for 128 μsecs to allow for the receiver to lock to its signal. It then sends a well-known Start of Frame Delimiter (SFD) to indicate the start of the physical payload. The first byte of the physical payload indicates the length (in bytes) of the payload itself. Its maximum value is 127, which limits the maximum length of a packet to 128 bytes. The receiver looks for a physical preamble to 'lock onto'. Once locked on, the receiver waits for the SFD and the length byte. It then estimates the earliest time after which it can switch off its radio.

2.2 Link Layer

In addition, to Low-Power PHY specifications, IEEE 802.15.4 also defines a MAC protocol. The specified MAC protocol proposes two network topologies, star and mesh. In star topology dedicated relay nodes referred as Full Function Devices (FFDs) are required to forward data from Reduced Function Devices (RFDs) to Network Coordinator (NC) node. On the other hand, in a mesh topology each node needs to act as a relay node as well. This mechanism is inappropriate in multi-hop network scenario, since the relay nodes need to maintain 100% duty cycle,

resulting in lower node and network life time. Similarly, the wireless channel is prone many degrading effects i.e. path loss, shadowing, interference, multi-path fading etc. Therefore, the proposed technique to use a single channel within a network results in link failure and instability in the network.

IEEE 802.15.4 MAC needs to be made compatible with the IoT requirements i.e. to enable high level robustness against wireless channel degradation effects, while maintaining a lower duty cycle preferably less than 1%. For this reason, IEEE 802.15.4e working group formed in 2008 [40], has been involved in addresses these issues. Dust Networks [34] introduced a channel hopping protocol named Time Synchronized Mesh Protocol (TSMP) [35] to mitigate multi-path fading which was later adopted by WirelessHART [36]. This TSMP inspired IEEE 802.15.4e working group to design a Time Synchronized Channel Hopping (TSCH) protocol, which later integrated into the IEEE 802.15.4 in 2010. In later part of this section we discuss the key functionalities of this TSCH MAC protocol.

2.2.1 Slot-frame Structure

In TSCH, a slot-frame is a group of slots which repeat over time and nodes get synchronize on this slot-frame structure. In a specific slot, a node can transmit, or receive, or sleep, according to a known schedule. In a sleeping slot, the radio is switched off by the node. However, in an active slot, the node transmits or receives to or from another node using a definite channel offset as specified by the scheduler on which channel offset. As shown in Fig. 2.2, a single slot is long enough to accommodate a complete maximum length packet transmission as well as an acknowledgment packet from the receiver side. The slot duration implementation-dependent; in [40] a value of 10 ms is suggested.

The MAC layer accepts packets from the higher layers and keeps them in its queue until they get successfully transmitted or discarded due to limit of retransmission count. Depending on the scheduler, a node decides whether to transmit the packet in the transmit queue or it needs to receive a packet, whatever the case the nodes stays in active state in these specific slots. Otherwise, if the transmit queue is empty and no packet is due from other nodes, the node switch to the sleeping state. IEEE802.15.4e also defines a simple backoff scheme to avoid collisions among channel competing nodes.

2.2.2 Scheduling

IEEE 802.15.4e defines the operation of the MAC under a given schedule; however it does not specifies how to plan a schedule. The schedule must be built considering mutual interference and availability of particular node(s) at definite slot time and the schedule should be updated periodically for reliable operation, since in wireless networks nodes association and disassociation is very dynamic.

In a centralized approach, a specific node referred as 'manager' is responsible

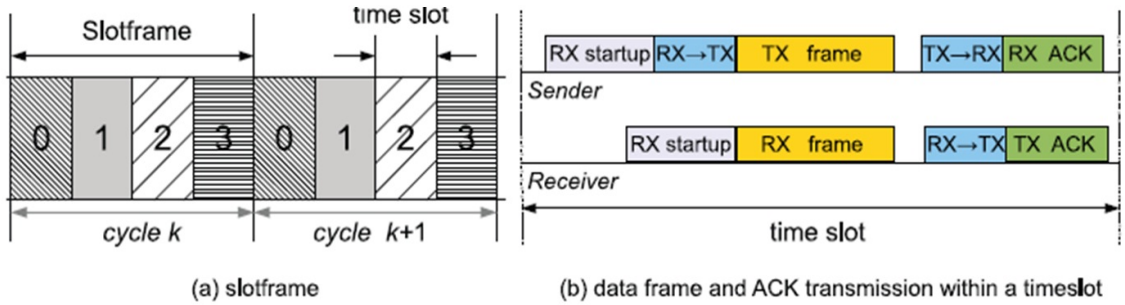


Figure 2.2: A 4-slot slot-frame and timeslot diagram of an acknowledged transmission.

for building and maintaining the network schedule. Each node reports to the manger about its neighbors and the amount of data it is generating. On the basis of this information, the manager draws connectivity graph and assign slots to each node depending upon its needs. The scheduler when shared with the nodes indicates them when they must be in active or sleeping state. For a less dynamic network scenario the centralized approach is very efficient as the topology is not much changing and thus effective centralized scheduling techniques can be employed.

In a highly dynamic network scenario where nodes come and go more frequently, a distributed approach is more effective, in which nodes decide locally on to schedule a link to each neighbor [41]. However, in a multi-path scenario when the data is generated at a constant rate then the scheduling of the intermediate links becomes a complex and challenging issue. In this case, Internet-like reservation protocols such as the Resource Reservation Protocol (RSVP) [42] and the Multi Protocol Label Switching (MPLS) protocol [43] could be applied, which still needs to be explored.

2.2.3 Synchronization

In a network based on slot-frame structure synchronization between devices is necessary. However, in TSCH (IEEE 802.15.4e) there is no beacon transmission, as shown in Fig. 1. Therefore, synchronization is achieved by either (i) Acknowledgment-Based synchronization (ABS) or (ii) Frame-Based synchronization (FBS). In ABS, the delta between the expected time of frame arrival and its actual arrival is shared with the transmitter by the receiver node in its acknowledgment. Thereby, transmitter gets synchronized to the receiver clock using this information. On the other hand, in FBS, the receiver node calculates the delta between the expected time of frame arrival and its actual arrival, and then the receiver node synchronized to the transmitter's clock. Whenever, there is communication between two or more nodes they resynchronize again and again. If two nodes have not communicated for some duration as long as 30 s, then these nodes

exchange keep-alive messages to resynchronized or identify link availability.

2.2.4 Channel Hopping

In addition to the time slot access, the IEEE 802.15.4e TSCH MAC promotes channel hopping. In [44] and [45], the performance gains achieved with help of channel hopping have been tested and analyzed. The wireless channels operating at different set of frequencies experience different level of interference and fading. Through frequency diversity that is achieved by channel hopping the effects of interference and multipath fading can be mitigated. In addition, the availability of multiple channels enabled multiple links to be active at the same time given the channels are orthogonal as in case of IEEE 802.15.4e, thus the capacity of the network increases. Moreover, combining time slot access with the channel hopping further improves reliability, since the probability of two transmissions at the same channel is minimized.

2.2.5 Network Formation

In a TSCH based network, the network is formed with the help of two using two components: advertising and joining. If a node intends to join the network, it listens for the advertisements, when from the neighboring nodes an Advertisement command frame is received; the new node sends a Join Request in response of the Advertisement command. If the network is centralized then this Join Request is forwarded to the Network Manager, which assigns a channel offset and/or a slot for transmission/reception to this node. Moreover, slot-frame is set up for the new node and draws links between this node and other neighboring nodes. However, in a distributed management system, the Join Requests are processed locally by the neighboring Advertizing node. These slot-frames and links can also be deleted and/or modified after a node has joined the network. Once the expected number of nodes has joined the network or there have been no Join Requests for some time, then the Advertizing commands can be periodically scheduled, so that the overhead associated with them can be reduced.

2.3 Network Layer

2.3.1 Internet Protocol

One of the most critical requirements of IoT is that the devices are IP enabled so that the communication between heterogeneous systems can be possible. Since, the number of devices connected in an IoT is going to be enormous, thus already exhausted IPv4 is no way a practical solution. However, it is not straight forward to enable IPv6 for low-power WPANs devices having limited resources in terms of energy supply, small packet sizes, low capacity, restrictive topologies,

highly unreliable communication, etc [46] - [47]. For example minimum MTU size for IPv6 is 1280 bytes, and without fragmentation it cannot be transmitted using 128 bytes maximum packet size of IEEE 802.15.4. Similarly, the 40 bytes IPv6 header also poses a significant overhead reducing bandwidth efficiency. To address these challenges, the IETF IPv6 over low-power WPAN (6LoWPAN) working group has defined an adaptation layer in RFC 4944 [48] and in RFC 6282 [49], whose main responsibility is to compress IPv6 headers and fragmentation of large packets. Similarly, other contributions from IETF working groups reduction in management and routing overhead, design of lightweight application protocol in [41], IPv6 addresses auto-configuration RFC in 2464 [50], subnet design recommendations in RFC 3819, the compliance with the recommendation on supporting link-layer subnet broadcast in shared networks [51], etc.

IPv6 Low-Power WPAN

The adaptation layer defined by IETF 6LoWPAN working group resides between IP and link layer. Its main responsibilities include suppression of redundant headers as some of the headers information can be inferred from other layers, secondly the fragmentation of the large sized packets to make them compatible with low-power devices [48], [49]. The presence or absence of a header depends upon the network needs; however the headers should be transmitted in a specified order. Depending upon their functionality, the headers are categorized in four types as given below:

- NO 6LoWPAN Header specifies that the received packet by the low-power device is not according to the specifications of 6LoWPAN, thus it has to be dropped. In this way, both 6LoWPAN and non-6LoWPAN compliant nodes coexist in a network.
- Mesh Addressing Header enables low-power devices to route packet through multi-hop links. This header leads other header of 6LoWPAN encapsulation. For each forwarder node, it includes the link-layer addresses of the considered forwarding node and of the next-hop node, in addition to the link-layer addresses of the Originator and of the Final Destination.
- Fragmentation Header is used when a datagram does not fit within a single IEEE 802.15.4 frame. This header includes: the Datagram Size, that is the dimension of the entire IP packet before link layer fragmentation; the Datagram Tag, which identifies the original fragmented IP packet; and the Datagram Offset specifying the offset of the (for 2nd and onwards) fragment from the beginning of the payload.
- Dispatch Header is used to compress an IPv6 header or to manage link-layer multicast/broadcast. All kind of Dispatch Header comes after Fragmentation Header except a Broadcast Header which comes prior to Fragmentation Header and it is used for controlled flooding or topology discovery. It

manages multicast/broadcast by using a 1 byte long Sequence Number for detecting and suppressing duplicate packets.

Header Compression

IPv6 headers carry lot of redundant information which can be useful in internetwork communications, however for intra WPAN communication this information can be inferred using other layers. For example, Payload Length can be estimated from the Datagram Size information if fragmentation is being used or from the MAC Frame Length. Similarly, Hop Limit can also be estimated by the transmitting node. The Internet Protocol Header Compression (IPHC) encoding technique [52] defined for LoWPAN can effectively compress both IPv4 and IPv6 packets. A 13-bit IPHC encoding field is attached with first 3 bits of the Dispatch Header. Depending upon the number of hops IPHC encoding scheme can compress unique local, global, and multicast IPv6 addresses up to 2 bytes (in case of single-hop). However, in case of multi-hop it can compress IPv6 headers up to 7 bytes. In addition to IPv6 addresses, 6LoWPAN enables to compress next headers using IPv6 Next Header Compression (IPNHC) encoding technique. The LoWPAN IPNHC field having variable number of bits follows IPHC field, can be used to decode next-headers in the same order they were encoded at the transmitter node. Similarly, LoWPAN NHC can be use to compress UDP headers, RFC 6282 [49]. Since, UDP Length field can be omitted estimating the size from MAC layer or fragmentation headers, transmitter and receiver ports may also be omitted if they are the same, in this way at least UDP header can be compressed up to 2 bytes, i.e. one byte for LoWPAN NHC field and another byte for compressed ports.

2.3.2 Routing Protocol

The battery operated low-power wireless nodes experience lossy radio links due to multi-path fading, topology changes due to mobility. Therefore, 6LoWPAN nodes require a robust routing protocol specialized for Low-Power Lossy Network (LLNs), considering IPv6 operation as well [53]. For this reason, the IETF working group Routing Over Low power and Lossy (ROLL) networks adopted a gradient-based approach [38], [54], [47], [55]. To design an effective IPv6 routing protocol for LLNs, RPL. This routing protocol RPL is compatible with multiple link layer technologies, specifically low-power based link layers, thus it is suitable applications in building/home automation, industrial environments, and other applications [56]-[57].

In a RPL based network scenario, some devices are opted to be root device and collect data by coordinating with other network device through multi-hop routes. Root device creates a Destination Oriented Directed Acyclic Graph (DODAG) for each other device based on the link costs and other device attributes. Then using an identifier DODAGID, this objective function is optimized for required conditions. Also, based on the distance between a device and its respective root device a

Rank metric is build which is used to set up a topology. The Rank monotonically decreases with the distance as per gradient-based approach i.e. farther the device from the root lowers the Rank and vice versa. The RPL protocol determines both the downward routes from root or gateway devices to end sensor or actuator devices for Point-to-Multipoint (P2MP) as well as the upward routes for the other way around which is used for Multipoint-to-Point (MP2P), these routes may also be used for Point-to-Point (P2P) communications as well.

The RPL topology can be as simple as network with only a single root device maintain a single DODAG, though this kind of implementation is feasible only for small number of network device. However, in case of a large network root device may maintain multiple virtual root devices responsible for only a part of the network. In this way, the DODAG becomes simple and flexible and the maintenance of the network can be done in efficient way. A more sophisticated and flexible configuration could contain a single DODAG with a virtual root that coordinates several LLN root nodes.

The information of end devices about the Rank, Objective Function, IDs, and so on, is embedded in DODAG Information Option (DIO) messages. These DIO messages are shared periodically by devices with their neighbors to create DODAG and route towards the root device. With the help of these DIOs devices compute their Ranks, switch between DODAGs as per implementation requirements identified by the objective function, and select a parent device among all the possible ones.

RPL routing protocol can exploit key network metrics such as: node energy, hop count, link throughput, latency, link reliability, and link color, to effectively adapt to the network channel conditions [58]. Specifically, RPL uses a metric (link color) to estimate based on the color of the link whether to include or exclude it for a particular DODAG. These metrics can be shared with other devices and the routes can be adaptively selected based on network conditions. However, more frequent changes results in instability in routes and degrade network performance.

2.4 Transport and Application Layer

The LoWPAN device with LLN when integrated with IPv6, can communicate through current world-wide Internet without the need of any Network Address Translation (NAT) gateways. Application layer protocols such as HTTP [12] provide semantics that facilitate content representation and inter-interoperability between client/server in a content/resource centric manner. The concept of IoT to enable IP connectivity for resource constraint devices is not realizable over the current higher layer communication protocols, since they are not designed for resource constraint low-power devices, therefore traditional communication protocols are not applicable for LoWPAN devices with LLN due to the extensive overhead and inefficient operation in a resource constrain scenario. For this reason, much work is going on to possibly suppress or at least compress the overheads of the

transport and application layer data assuring same level of application experience as enjoyed the current Internet.

2.4.1 Transport Layer

In IP based networks end-to-end reliability is provided by the Transport layer [14]. In addition, using Automatic Repeat- Request (ARQ) [59], TCP also provides flow control and congestion control. To provide end-to-end reliability additional overhead is required which only reduces the bandwidth efficiency as well as the end-to-end packet reception acknowledgments consume significant amount of energy. On the other hand, User Datagram Protocol (UDP) [60] provides best effort transport with no guarantees of packet delivery. However, control the retransmission mechanism at the application layer and less costly UDP in terms of overhead, can give an effective tradeoff between energy efficiency and reliability, which can be optimized based on the application and implementation requirements.

UDP is already a lighter transport protocol as compared to TCP; however UDP headers can be further compressed by using LoWPAN NHC [49]. Since, UDP Length field can be omitted estimating the size from MAC layer or fragmentation headers, transmitter and receiver ports may also be omitted if they are the same, in this way at least UDP header can be compressed up to 2 bytes, i.e. one byte for LoWPAN NHC field and another byte for compressed ports. Comparatively, UDP is considered to be more feasible for IoT based on 6LoWPAN architecture over TCP.

2.4.2 Application Layer

The application layer protocol strictly depends upon the network scenario and the application and implementation requirements of the system. Not all kind of application layer protocols suit to all kind of devices, for example in case of resource constraint low-power devices in a LLNs, cannot be operated using traditional client/server model defined by HTTP. While REST architectures offer reliable communication without considering the pay load size, which may require fragmentation in IEEE 802.15.4 based device and as the fragmentation requires additional overhead that is why the operation may become practically infeasible. Therefore, limiting the packet size at the application layer is a necessary condition to make the operation compliant with restrictive lower layers.

The IETF Constrained RESTful Environments (CORE) working group [61] proposed the Constrained Application Protocol (CoAP) [39]. The CoAP translates to HTTP making it compliant with resource constraint devices; this is achieved by multicast support, reduced overhead, and simplicity. CORE working group designed CoAP specifically designed for LLNs [62] and as a subset of RESTful architecture, thus offering compatibility with the HTTP [12]. The main features addressed by CoAP [39] are:

- Constrained web protocol specialized to M2M requirements.
- Stateless HTTP mapping through the use of proxies or direct mapping of HTTP interfaces to CoAP.
- UDP transport with application layer reliable Unicast and best-effort multicast support.
- Asynchronous message exchanges.
- Low header overhead and parsing complexity.
- URI and Content-type support.
- Simple proxy and caching capabilities.
- Optional resource discovery.

In the HTTP client/server based architecture the end points have definite roles. However, CoAP is an asynchronous request/response protocol based on a datagram oriented transport i.e. UDP, where both endpoints act as clients and servers. The architecture of CoAP is divided in two layers, a message layer and a request/response layer.

1. Message layer: The function of the CoAP message layer is to control message exchanges over UDP between two endpoints. Requests and Responses share a common message format. Messages are identified by an ID used to detect duplicates and for reliability.
2. Request/Response layer: CoAP request and response semantics are carried in CoAP messages, which include either a method code or response code, respectively. Optional (or default) request and response information, such as the URI and payload content-type are carried as CoAP options. A Token Option is used to match responses to requests independently from the underlying messages. As CoAP is implemented over non-reliable transport, CoAP messages may arrive out of order, appear duplicated or be lost without notice. Thus, CoAP needs to implement a reliability mechanism with the following features:
    - Simple stop-and-wait retransmission reliability with exponential backoff for confirmable messages.
    - Duplicate detection for both confirmable and non-confirmable messages.
    - Multicast support.

2.5 Infeasibility of Current IoT Protocol Stack for IoM

The present protocol stack for IoT is defined considering tiny low-power objects or things those are communicating and exchanging scalar data or periodically acquired physical measurements from the environment. However, the nature of multimedia data specifically real-time continuous multimedia data is burst. Since, the nature of data in IoM and communication requirements of IoM have been over looked in present IoT protocol stack, that is why present communication mechanisms designed for IoT are not adequate for IoM traffic requirements. In the following part of this section, we discuss the limitations of the currently proposed IoT communication protocols in support of IoM with respect to energy efficiency, higher data rate and Internet connectivity.

The IEEE 802.15.4 [32] defined PHY and MAC layer specifications give a well suited trade-off between energy-efficiency, transmission range, and data rate, which makes it most suitable for small to medium sized WPAN. The maximum supportable data rate and maximum packet length of 250 kbps and 128 Bytes, respectively, has the capability to perform well in many IoT applications in which only a limited and/or periodic data is communicated between objects or things. However, the multimedia bandwidth requirements are significantly higher (as shown in Table 5) as compared to the data rate offered by IEEE 802.15.4 i.e. 250 kbps. Similarly, in case of real-time continuous multimedia communication the continuously generated data requires large packet size, so that the bandwidth is utilized efficiently maintaining the delay and jitter under the required thresholds. However, the IEEE 802.15.4 limited maximum packet size 128 Bytes increases delay and jitter, as well as it requires fragmentation at higher layer which comes with additional protocol overhead and increases the cost in terms of bandwidth and energy efficiency.

The MAC layer Time Synchronized Channel Hopping (TSCH) protocol proposed in the revised version of IEEE 802.15.4 (in 2006), efficiently exploits multiple orthogonal channels to allow multiple active links simultaneously. When combined with slot-frame structure, this approach allows nodes to maintain lower duty cycle as low as 1 %. The time slot structure can be inefficient in terms of bandwidth utilization, when nodes are variable traffic requirements. This will result in some slots to go idle even when some other nodes need strict bandwidth requirements. Therefore, a MAC protocol for IoM is required which provides specific QoS depending upon the traffic requirements of the node. In addition, due to the continuous traffic nature of multimedia applications, maintaining a very low duty cycle is hard providing the delay or jitter constraints. Thus, additional energy saving techniques should be adopted such as high data rate enabled MAC layer which can transmit larger amount of data for the same duty cycle size as compared to the currently employed IEEE 802.15.4 data rate i.e. 250 kbps.

The communication links between low-power devices are lossy, that is why a routing protocol (i.e. RPL) is proposed for resource constraint devices forming

Audio Source	Sampling Rate	Bits/Sample	Bit Rate
Telephone Grade Voice (up to 3.4 KHz)	8000 samples/sec	12	96 Kbps
Wideband Speech (up to 7KHz)	1600 samples/sec	14	224 Kbps
Wideband Audio Two Channels (up to 20 KHz)	44.1 Ksamples/sec	16 per channel	1.412 Mbps for both channels
Image Source	Pixels	Bits/Pixel	Bit rate
Color Image	512 x 512	24	6.3 Mbps
CCIR TV	720 x 576 x 30	24	300 Mbps
HDTV	1280 x 720 x 60	24	1.327 Gbps

Figure 2.3: Multimedia Applications and their Bandwidth requirements.

the IoT which is specially designed for low-power and lossy links networks. The RPL is a proactive routing protocol that is the routes are determined between every two device in the network not matter the two devices want to communicate or not. The routes are maintained using DIO messages which are transmitted by each device in the network and the status of an already build link is known through keep alive messages. It seems that the availability of the routes can be supportive in providing lower latency or jitter for multimedia traffic. However, it has been investigated in [63] that the proactive routing protocols have sub-optimal performance in multimedia communication in terms of packet end-to-end delay, overhead per packet, packet delivery ratio. It is due to the unnecessary route establishment and maintenance overhead which wastes bandwidth and energy. Thus, the current routing protocol designed for IoT based on resource constraint devices is infeasible for multimedia traffic and a low overhead preferably a reactive routing protocol is needed which considers network metrics to establish links like RPL, yet requires less routing overhead.

The TCP [14] provides reliable transport layer communication at the cost of increased communication overhead. Since, MAC layer also offers measures for reliable communication that is why for 6LoWPANs instead of TCP, the low overhead UDP [60] is promoted. In addition, further overhead reduction is achieved using header compression techniques. As far as the multimedia traffic is concerned, UDP has also ready been accepted as the more suitable transport layer protocol particularly for real-time multimedia communication. However, to improve multimedia communication other transport layer protocols designed on top of UDP could also be used such as the Lightweight User Datagram Protocol (UDP Lite) [64] which provides data as well as partial checksum (only dropping the packet if header is erroneous) or the Datagram Congestion Control Protocol (DCCP) [65] which has a little larger header size (12 to 16 Bytes) as compared to UDP (8 Bytes), nevertheless it is a connection oriented like TCP and provides path MTU, data checksum, partial checksum, and congestion control which makes it more suitable for timing constraint multimedia application.

From the discussion above it can be concluded that the currently employed communication stack for IoT is in feasible for multimedia communication in 6LoWPANs in terms of bandwidth utilization and energy efficiency. Thus, to realize IoM based on resource constraint devices, energy aware and bandwidth efficient protocols are required to facilitate multimedia communication considering acceptable level of jitter, end-to-end delay, latency, etc over low-power and lossy links networks.

Chapter 3

IEEE 802.11 for Internet of Multimedia

3.1 Introduction to IEEE 802.11

IEEE 802.11 defines media access control (MAC) and physical layer (PHY) specifications for implementing wireless local area network (WLAN) in the 2.4, 3.6, 5 and 60 GHz frequency bands. IEEE 802.11 WLANs provide a maximum data rates of 11 Mbps up to 1 Gbps, depending upon the standard employed. WLANs have been widely deployed in public and private areas in recent years. Meanwhile, more and more portable and mobile devices, such as mobile phones, PDAs, laptops, are equipped with WLAN interfaces, allowing users to access mobile Internet applications and services via WLANs [66].

In the IEEE 802.11 protocol [67, 68], the fundamental wireless random medium access mechanism is Distributed Coordination Function (DCF), which employs a contention-based MAC protocol, called carrier sense multiple access with collision avoidance (CSMA/CA) [68, 69, 70]. There are two operating modes used for IEEE 802.11 based WLANs, (i) Infrastructure Mode is used when there is at least one Wireless Access Point and client. The client connects to the network through the Access Point to gain internet access. (ii) Ad Hoc Mode is used when wireless clients want to directly communicate with each other without going through an Access Point. This is also called peer-to-peer mode.

3.2 Feasibility of IEEE 802.11 for IoM

ZigBee that is based on the IEEE 802.15.4 standard defines specifications for WPANs formed by simple devices that are low-cost and low-power devices consuming minimal amount of energy. It provides self-organized, multi-hop, and reliable mesh networking with long battery lifetime [71]. Although Internet Protocol (IP)

IEEE 802.15.4 with 6LoWPAN (IPv6 over Low-Power WPAN) adaptation layer. ZigBee standard is designed considering that the network devices perform simple operations and thus require a lower data rate that is why ZigBee only supports a maximum of 250 kbps. However, due to the popularity of multimedia application motivating towards the IoM, this supportable data rate is much lower. Especially for real-time multimedia communication, the multimedia devices cannot provide satisfactory user experience with IEEE 802.15.4 standard. For this reason, IEEE 802.11 standard is suggested in the literature for wireless multimedia network, since it provides a high data rate communication model and holds a great potential for wireless multimedia sensor network.

ZigBee and other IEEE 802.15.4 based protocols have been considered for WSNs applications due to their energy-efficient design. However, recently developed power-efficient Wi-Fi models promise multiple years of battery lifetime, have become a strong candidate in this domain [23]. Reuse of existing Wi-Fi infrastructure offers cost savings and faster deployments. Widely deployed Wi-Fi networks reduce the infrastructure cost to a minimum while improving the total cost of ownership. Wi-Fi devices have the advantage of native IP-network compatibility, which is a big plus for IoT. Well-defined and universally accepted IP connectivity overcomes the need of expensive gateway requirements or any network address translation (NAT). The communicating devices can be assigned static or dynamic IP addresses using Dynamic Host Configuration Protocol (DHCP) servers and to resolve network addresses to link layer addresses the Address Resolution Protocol (ARP) can also be used.

Apart from the IP or Internet connectivity support, IEEE 802.11 provides various mechanisms to support different types of traffic classes with desired Quality of Service. Similarly, to provide reliable communication Wi-Fi Protected Access (WPA/WPA2) provides link layer encryption and authentication support. the random access mechanism and higher data rate support provides support to improve latency, end-to-end delay, and jitter etc, in other words provides better user experience for multimedia communication. In addition, IEEE 802.11 provides reliable communication using acknowledgments and transmitted packets ordering support. Moreover, it can support transmission range corresponding to the view range of video devices that makes it a practical choice. Furthermore economy of scale is another important advantage of Wi-Fi with an expected 22 percent annual growth rate between 2010 and 2015 [24]. For these reasons, Wi-Fi has already been widely accepted for many commercial off-the-shelf video devices, which are largely deployed for video surveillance and monitoring applications, making it a good candidate for IoM if the energy efficiency mechanisms of IEEE 802.11 are devised comparable to IEEE 802.15.4.

3.3 Power Saving Mechanisms Defined in IEEE 802.11 Standards

3.3.1 IEEE 802.11 a/b/g

First standard for IEEE 802.11 WLANs appeared in 1997, which defines the WLAN MAC and PHY Layers specifications that provide data transmission rates up to 1 Mbps and 2 Mbps over the 2.4 GHz range, which was further upgraded to 11/54/54 Mbps in IEEE 802.11 a/b/g standards in 1999/1999/2003 respectively. All these standards adopted two medium access mechanisms DCF (default) and PCF (optional). In DCF, no power management mechanism was employed to save power. However, using RTS/CTS packets virtual carrier sensing is proposed to defer transmission in the transmission range of the active transmitter and receiver, to avoid collisions. Deferring stations set Network Allocation Vector (NAV) and go to the doze state which consumes minimal energy, for a duration indicated in RTS/CTS packets. However, in PCF two modes of operations for wireless stations were defined; active mode and power saving (PS) mode. In active mode a station remains fully powered on and able to receive/transmit or sense channel activity. Whereas, in PS-mode a station wakes up only to receive beacons after some listening interval this can be a multiple of beacon intervals [72]. If there is a traffic indication map (TIM) indicating that there is a packet buffered for this particular station then it switch to the active mode. In PCF a beacon interval is divided into two periods; contention free (CF) period in which stations operating in PS-mode send or receive data, and a contention period in which all stations PS or Non-PS stations operate using DCF. In infrastructure mode buffered broadcast/multicast and Unicast packets are indicated using simple TIM or Delivery-TIM frames. If there is a packet for a station indicated in TIM, it sends a CF-poll request frame to AP, who responds with the packet and station sends acknowledgment. In ad hoc mode, first stations get synchronized and then using announcement-TIM stations indicate other stations about their future frame exchanges. If a station doesn't has a packet to send or receive it switches to the doze state for this beacon interval. Only significant difference between DCF and PCF is that in PCF there is no backoff period and no collision period, since AP coordinates all channel accesses.

3.3.2 IEEE 802.11 e

DCF is based on the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol. However, it is unable to provide the required performance for voice and video applications, because it is fundamentally developed for Best Effort services [73]. For that reason, the IEEE 802.11e 2005 amendment was approved in order to provide QoS support to WLANs. The EDCA and HCCA schemes are enhanced MAC access mechanisms which are proposed in IEEE 802.11e [73]. The EDCA scheme improves the DCF by assigning various

sizes of IFSs and contention windows (CWs) to packets associated with different access classes. Therefore, the EDCA scheme makes it possible to achieve per-class QoS. Unlike DCF, in EDCA the periods for overhearing the transmissions between other STAs and the AP and the waiting backoff periods vary (different power consumption) for packets belonging to different access classes. Whereas, the HCCA scheme uses a hybrid coordinator (HC), as a centralized coordinator to allocate a time period, called the controlled access phase (CAP). The HCCA scheme can fully manage radio resources during CAPs and grants transmission opportunities (TXOP) to STAs for sending back-to-back packets without contention. The power consumption of the HCCA is comparable to the PCF.

The Automatic power save delivery (APSD) defined in IEEE 802.11e suggests two mechanisms, i.e., the Scheduled APSD (S-APSD) and Unscheduled-APSD (UAPSD). In APSD mechanism service period (SP) is assigned to stations for exchanging packets with the AP. Since, the station does not have to contend the channels, thus the power consumption reduced. The S-APSD considers the characteristic of packets which are generated periodically; thus the AP allocates SPs to the station periodically. Then, the station only wakes up periodically, receives and sends packets with the minimal contentions. On the other hand, the AP does not offer SPs periodically to a station in the U-APSD. In the U-APSD, the station can send an uplink frame to trigger an unscheduled SP for exchanging the packets with the AP. Although, the station has to contend the channel for sending the up-link trigger frame, the U-APSD method improves the IEEE 802.11 PSM by averting the PS-Poll procedure, and save more station energy. For example, an uplink voice packet can be configured as a frame to trigger a service period, which is used to transmit downlink voice packets. Nevertheless, these mechanisms have the potential to offer significant power savings to the clients. However, the benefits depend on the exact usage scenario, such as the number and type of concurrent clients as shown in [74].

3.3.3 IEEE 802.11 n

Face to demand for higher performance WLANs to support multimedia applications, the standard IEEE 802.11n is appeared for next generation WLAN [68, 20, 75]. An IEEE 802.11n WLAN can operate with transmission data rate reaching 600 Mbps [68, 76, 77]. Three main MAC enhancements have been appeared with 802.11n to reduce the protocol overheads are: Aggregation MAC Service Data Unit (A-MSDU) and MAC Protocol Data Unit (AMPDU) and Block Acknowledgment (BA) [68, 20, 78, 79]. The principle of MSDU aggregation is to allow multiple MSDUs to be sent to the same receiver concatenated in a single MPDU, whereas the principle of MPDU aggregation is to join multiple MPDUs to be sent with a single PHY header. Since, the much protocol overhead is reduced thanks to frame aggregation and block acknowledgment, therefore a lot of energy is saved that could have been utilized by protocol overhead otherwise.

In a WLAN, clients can exploit at least two kinds of sleep opportunities. First, when an AP is transmitting to client A, another client B can be put to sleep like virtual carrier sensing. Second, interframe gaps can be exploited by putting the client to sleep for the duration of the gap. This is, of course, idealized, since it assumes that interframe gaps are known before-hand, and that the client can be put to sleep for arbitrarily small intervals (a capability sometimes called micro-sleep).

In IEEE 802.11n predecessor standards, an AP broadcasts a TIM frame indicating the STAs to receive queued packets. An STA which receives a TIM frame must contend for the channel, and send a PS-Poll frame to the AP. However, PS-Poll frames sent from multiple stations to the AP at the same time can collide and stations have to contend again for PS-Poll frame transmission. The WLAN resources are wasted, and the STAs also consume extra energy. To avoid this, the power save multi-poll (PSMP) scheme is defined in IEEE 802.11n. The AP could consider different QoS requirements, such as delay constraints and bandwidth constraints, and schedule the packet transmission of stations using the multi-polling mechanism. The AP specifies the schedule information in the beacon frame so that the STAs can wake up and receive packets based on the AP schedule. The PSMP mechanism improves both energy efficiency of PSM stations and WLAN utilization by minimizing PS-Poll contentions.

Unlike its ancestors, the 802.11n standard enables multi-input-multi-output (MIMO) transmission and reception capabilities, where multiple antennas can be simultaneously used to increase throughput or provide diversity. Accordingly, 802.11n NICs have an antenna selection capability in which a subset of antennas can be chosen for transmission or reception. The RF-processing components (or RF-chains) corresponding to unused antennas can then be powered down. This creates additional device power states, each with different power consumption and Idle, TX and RX energy costs vary nonlinearly with the number of RF-chains used. Moreover, sleep mode power consumption is dramatically lower than other modes.

3.3.4 IEEE 802.11 ac

Even though 802.11n can deliver high throughput, only one-to-one communication in the infrastructure mode is supported and the network throughput is limited by the maximum per link data rate [80]. To overcome this deficiency, a new standard IEEE 802.11ac [81][82][83] is under development adding enhancements for PHY and MAC layers. The main PHY layer enhancements are: the 80MHz bonding technique, 8*8 MIMO antenna support, and Downlink Multi-User MMIO (DL MUMIMO) which is used to allow multiple frames to be sent from the AP to multiple receivers simultaneously through multiple spatial streams. The main MAC layer enhancement is TXOP sharing [84] which is used to perform multiple downlink traffic streams to multiple receiver STAs simultaneously. Another MAC layer enhancement is the extension of the maximum aggregation size for both

A-MSDU and A-MSPU.

3.4 IEEE 802.11 MAC Functionalities Exploitable to Save Energy

In order to optimize the current IEEE 802.11 MAC for the adaptation in IoM, following parameters are identified which can be tuned to achieve the desired performance.

3.4.1 Reducing contention overhead by tuning contention window size

IEEE 802.11 access mechanism is based on wireless random channel access. Thus, stations defer their transmissions by taking a binary exponential backoff. Contention duration depends upon the minimum contention window (CWmin) size as well as to the number of stations in a BSS. If a large CWmin size is selected then the probability of packet collisions is reduced for the same number of stations. Thus, the retransmissions per packet are decreased, saving energy. However, due to the higher CW size much of the time is spent in idle listening which itself consumes energy, decrease effective throughput and increases delays. Decreasing the CW size increases the number of collisions and thus increasing number of retransmissions (and end-to-end delay) and significant energy is utilized in retransmissions. To efficiently decrease the contention overhead, the CWmin size should be selected considering the tolerance level of the delay per packet and the number of stations within the interference range. For example, for a sparse network scenario (less number of stations) the CWmin size should be decreased such that the end-to-end delay per packet remains under the tolerable level.

Generally, the power saving algorithms perform poor in terms of per packet delay, therefore in [85] a backoff technique is proposed which reduces energy consumption in contention considering quality-of-service in terms of delay. The proposed scheme compels wireless stations to stay in power saving state during backoff period and do not listen on the channel for other transmissions. In this way, backoff counter never freeze as a result delay is reduced, however after backoff expires channel is sensed as if it is idle in order to avoid collisions. To reduce the number of per packet retransmissions due to the collisions, a novel mechanism to select optimal CWmin and CWmax size is proposed in [86]. The proposed mechanism selects optimal CW sizes based on the number of 1-hop neighbors and the remaining amount of energy as well as the number of retransmission attempts per packet. In both [87] and [88], using an analytical model it is shown that not only the size of the CW affects the energy efficiency of WLANs but the packet size also significantly effects energy efficiency. Specially in harsh channel conditions (high BER),

the packet size becomes crucial as the energy efficiency increases with the packet size for some level then drops considerable with any further increase in packet size.

3.4.2 Reduce overhearing of the transmissions destined for other stations

In current IEEE 802.11 architecture, all stations listening to the channel, receive the packet that is being transmitted. After receiving the packet is sent from the physical layer to the MAC layer that checks as if the packet is destined to it or not. If the packet is not destined for it, MAC layer drops the packet. This process results in wastage of energy, since reception of a packet consumes significant amount of energy. Thus, the unicast and multicast packets should be scheduled or transmitted in such a manner that reduces the overhead of overhearing the transmissions.

In [89] formulas are derived for the amount of energy that a station consumes to transmit 1 MB of data in an IEEE 802.11 network with n stations. It is reported that the useful energy consumed in successful transmission and reception of data is constant; however significant amount of the remaining energy is wasted due to overhearing, listening to an idle channel, unsuccessful (colliding) transmissions, and reception of collisions. Through RTS/CTS mechanism significant amount of energy is saved with the help of NAV. However, present CSMA/CA mechanism doesn't permits to reduce idle listening. In [90] a time window is detected in which a station overhears the packet transmissions destined for other nodes. Therefore, during this time window the wireless stations switches to the low power energy saving idle state. Since, idle state consumes less energy as compared to the receiving state (while overhearing), thus significant amount of energy is saved.

3.4.3 Avoid packet losses, or speed up transmission

The packets in IEEE 802.11 based WLANs can be lost due to two reasons, either the packet transmission is collided with some other packet transmission at the receiver or the packet is lost due to wireless channel induced errors as a result the received packet is unable to be decoded correctly. No matter what was cause, the packet has to be retransmitted which consumes significant energy, since the whole packet transmission process has to be repeated involving higher contention due to BEB, contend for channel access and retransmission of the whole packet. However, in a scenario where packets are lost due to collisions then the transmission rate (data rate) should be increases so that the overhead of the transmission can be reduced. Otherwise, if the packets are lost due to the bit error rate (BER), then the transmission rate should be decrease since high data rate modulation and coding schemes are not robust against high BER. By using these adaptive transmission rates control techniques, a significant amount of energy can be saved that is consumed in retransmissions.

To reduce the number of per packet retransmissions due to the collisions, a

novel mechanism to select optimal CWmin and CWmax size is proposed in [86]. The proposed mechanism selects optimal CW sizes based on the number of 1-hop neighbors and the remaining amount of energy as well as the number of retransmission attempts per packet. It is proposed in [91, 92] to reduce transmission rate upon severe packet losses. However, these approaches perform poor in hidden node scenario, where packet is being lost due to collisions and not due to channel conditions. Reducing the transmission rate in this scenario increases the packet transmission time which further increasing the probability of collisions. Similarly, reducing or increasing the transmission rate upon failure or success in consecutive attempts, respectively, has been promoted in [91, 92]. However, in [93] it has been proved that in practical scenarios these predictive mechanisms are infeasible, due to the random behavior in packet losses.

3.4.4 Controlling the transmission power

The amount of energy consumed by a station and its operation is a network critically depends upon the level of the transmission power being used. For example, if the transmission power of the access point (AP) is set to a high level, then using less number of APs coverage can be provided to a larger area. However, only the stations close to the AP can enjoy high data rates and the stations on the edges take much of the air time due to their lower data rates. In addition, increasing the transmission power also increases the interference range. Thus, the transmission power should be increased when there is less number of stations in the network and there is no higher data rate requirement. On the other hand, if higher capacity is required, then the number of APs should be increased for the same coverage area. Thereby, each AP has a small region to cover, thus its transmission power can be decreased. In this way, interference among stations reduces and end-to-end delay as well as throughput of stations improves.

The power level for transmitting a packet can be increased or decreased to get higher coverage or reducing interference with other flows, respectively. Generally, the transmit power at all stations is assumed to be same (symmetric). However, if the transmission power of per packet is tuned for a transmitting station depending upon the distance and the channel conditions at the receiving station, then probability of transmitting a packet successfully can be increased minimizing interference with other stations or links. This adaptive per link power control mechanism has been shown in the literature to improve energy efficiency and network throughput, especially in ad hoc network scenario.

An energy efficient adaptive transmit power and transmission rate control mechanism is proposed in [94]. However, the proposed mechanism only considers energy consumption at the transmitter side, whereas in [95] it is reported that overall network energy critically depends upon the energy consumption at the receiver. Therefore, in [96] an energy efficient mechanism is proposed which considers energy consumption at both transmitter and receiver end, such that

overhearing at other nodes could be minimized, thus reducing the energy wastage at the non-destined receiving stations.

3.4.5 Reduce the transmission time

During the transmission state, a WLAN station consumes the maximum level of energy as compared to other states i.e. receiving, idle, sleep. To decrease the energy consumed during this state, the transmission rate of the station should be increase so that the transmission time could be reduced. However, higher data rates are not robust against BER, therefore retransmissions per packet may increase in poor channel conditions. Although the retransmission overhead will also be lower since per transmission time is lower due to higher date rate yet the transmission rate should be maximized up to a level which ensures specific end-to-end delay requirements and overall per packet transmission time. In a network scenario where BER is not significant and packets are lost due to multiple simultaneous transmissions (collisions), then increasing the transmission rate will reduce end-to-end delay as well as the retransmission overhead as a result the less energy will be consumed for per packet transmission.

It is reported in [95], that the energy consumption at the receiving nodes is crucial for the overall network energy efficiency. In addition, it is also shown that for making the transmitting station more energy efficient the transmission rate should be decreased, however to make both transmitting and receiving stations energy efficient, higher transmission rates are more suitable. In other words, modifying the transmission rate alters the transmission time for a packet. To reduce the probability of collisions due hidden terminal scenarios, in [93] it is suggested that the transmission time should be decreased by increasing the transmission rate, so that the cost of collision is reduced as well as the contention time.

3.4.6 Link adaptation or Adaptive modulation and coding

In link adaptation process, channel state information (CSI) at the receiver is shared with the transmitter, which depicts the combined effect of path loss, fading, interference etc at the receiver. Thereby, the transmitter adapts a suitable modulation and coding scheme or other protocol parameters like transmit power etc so that the packet can be successfully transmitted using highest possible data rate under the given channel conditions. The process reduces the number of retransmissions per packet and improves effective throughput which results in saving significant amount of energy.

It is proposed in [91, 92] to reduce transmission rate upon severe packet losses. However, these approaches perform poor in hidden node scenario, where packet is being lost due to collisions and not due to channel conditions. Reducing the transmission rate in this scenario increases the packet transmission time which

further increasing the probability of collisions. Similarly, reducing or increasing the transmission rate upon failure or success in consecutive attempts, respectively, has been promoted in [91, 92]. However, in [93] it has been proved that in practical scenarios these predictive mechanisms are infeasible, due to the random behavior in packet losses.

3.4.7 Tuning Fragmentation Threshold to reduce retransmissions

Large sized packets transmitted over a WLAN channel are more prone to get erroneous due to various wireless channel effects like path loss, interference, multipath fading etc as compared to smaller packets. The 802.11 standard includes the ability to fragment Unicast packets for improving throughput performance in the presence of poor channel conditions. The transmitting station divides the incoming higher layer packet into smaller fragments if it is larger than a predefined fragmentation threshold level (256-1024 B). These fragments are sent separately to the receiving station. Each fragment consists of a MAC Layer header, frame check sequence (FCS), and a fragment number indicating its ordered position within the frame. FCS indicates if more fragments to be sent or not. Since, the source station transmits each fragment independently, the receiving station replies with a separate acknowledgment for each fragment.

Fragmentation comes with additional protocol overhead as opposed to frame aggregation mechanism. However, if the WLANs are experiencing poor channel conditions and/or the number of per packet collisions is high then fragmentation can improve the overall network throughput and save energy that is utilized in retransmissions. Similarly, a large packet is dropped at the receiver even if only a part of the frame get erroneous, due to frequency selective fading, yet the whole packet is transmitted. Whereas, in fragmentation process if any fragments are lost. Then only these particular fragments are retransmitted and not the whole packet. Therefore, the energy utilization per packet is reduced if the fragmentation level is chosen appropriately.

An adaptive mechanism is proposed in [97] for varying the MAC layer frame sizes in order to partition the network-bound data, to enhance throughput and energy savings. This is achieved by computing a frame length based on the current medium conditions characterized in terms of the BER. An approach that dynamically adapts fragmentation sizes, retransmission retry limit, and transmission power for achieving energy efficiency in wireless systems was presented in [98]. However the optimal fragment sizes computed in these efforts do not consider the latency and energy costs incurred during the fragmentation process. Therefore, in [99] present a novel feedback-based mechanism that dynamically adapts the fragment sizes based on the observed end-to-end latencies and channel conditions to achieve energy efficiency while satisfying the real-time constraints of the application.

3.4.8 Reducing Interframe Space and Shorter preamble

Reduced Interframe Space (RIFS) is 2 sec time duration used in place of Short Interframe Space (SIFS) which is typically 16 to 20 sec in block acknowledgment and/or in frame aggregation. It is used between a MSDU and a block-Ack request frame and/or between back-to-back MSDUs in IEEE 802.11n WLANs. Since, its duration is notably lower than SIFS, thus it improves throughput and energy utilization.

Preamble is used to inform the receiver that a data packet is coming through and for synchronization purpose. Short preambles were introduced in IEEE 802.11 b radios but its operation was optional; however in IEEE 802.11 g it was mandatory. The short preamble is composed of 72 bits and long preamble is 72 bits long. Clearly, short preamble will require lesser time and energy to get transmitted so short preamble is preferred where applicable.

3.4.9 Adaptive Retry limit

The IEEE 802.11 standard allows the MAC to drop the higher layer packet if it is not be successfully transmitted in some given number of transmission attempts, this number is referred as retry limit. However, the decision to drop a packet depends upon the application. For example, for a real time multimedia application instead of retransmitting based on retry limit, the packets should be dropped if it has been in the MAC layer queue for a specific duration such that even if the packet is successfully get transmitted after this time the required end-to-end delay requirement is not fulfilled. Therefore per packet based retry limit should be used instead of symmetric retry limit for all WLAN stations. Thereby, in a multi-hop network scenario the packets should be incorporated with the time they have been generated at the source, so that the intermediate stations may utilize this information combined with the estimated duration required to further process the packet, in order to drop the packet. In this way, a lot of energy can be saved by reducing the useless retransmission attempts.

Chapter 4

Energy Saving Transmisison Strategies for IEEE 802.11

4.1 Strategies at the PHY Layer

We have proposed the use of MIMO-OFDM based 802.11n physical layer to be used in the designed camera nodes. A lot work in the literature address the power consumption issue for Wi-Fi [100] [101] [102] [103] [104] [105]. The work can coarsely be divided into two categories. The first work is the development of statistical power consumption model based on certain traffic models and network performance parameters e.g. in [106]. In the other category, most of the literature works ground their analysis in experimental measurements. In the latter category either the total energy consumption is physically measured at the wireless network interface [107] [1] [108] or experiments are designed to decompose the energy consumption of the wireless interface into its various constituents.

The camera nodes in the Internet of Multimedia Things have to transmit continuous streams of visual information. The transmission rate varies depending upon the captured scene. Given a sizeable configuration space at the physical layer of 802.11n, the search is for optimal configuration which minimizes the energy consumption while maintaining the requirement video quality.

In [1], the power consumption statistics for a wireless network interface (WNI) operating over a broad set of configurations, are reported. In [107], similar statistics for Google Android G1, Magic, Hero and Nexus handset have been reported. The power consumption is decomposed into its constituents in [2], where the energy while passing the data across various layers, to the kernel, and finally to the ASIC have been distinguished. In this text, we focus on the results presented in [1], [107] and [2].

In 4.2 the configuration space for the 802.11n standard is discussed in detail, followed by the detailed analysis of power consumption for various configuration in 4.3.

4.2 Tuneable Parameters

802.11n is a wireless LAN standard, which was finalized in 2009. It belongs to the family of IEEE's wireless fidelity (WiFi) standards. 802.11n is based on the MIMO-OFDM technology. The standard specifies the transmitter specifications and barring a few aspects, generally leaves the receiver design to the implementer.

4.2.1 Transmitter Overview

The transmitter block is fed a stream of bits. These bits could be the binary representation of some file stored in memory or could also be generated randomly. The transmitter process chain beings with forward Error Correction (FEC), which is applied to the incoming bit sequence to increase reliability of transmission. Stream parser divides the encoded bits into different streams that are sent to different interleaver and mapping devices. The sequence of the bits sent to an interleaver is called a spatial stream. Interleaver changes order of bits for each spatial stream to prevent long sequences of adjacent noisy bits from entering the FEC decoder. FEC scheme implemented in this quarter is Low Density Parity Check Code (LDPC) for which interleaving is not used as mentioned in standard. Constellation mapper maps the sequence of bits in each spatial stream to constellation points (complex numbers).

Afterwards, the modulated data symbols along with guard band and pilot subcarriers are assigned to sub-carriers in a particular order. IFFT is applied to the entire packet to generate a time domain data sequence. This is followed by the insertion of cyclic prefix at the start of each OFDM symbol to provide robustness against the delay spread of the wireless channel, thus avoiding Inter Symbol Interference (ISI). Finally transmit sequence for each spatial stream is space-time coded, to further increase its robustness against the channel impairments. Data streams coming out of space time encoder are called space time streams. Spatial mapper maps space time streams to transmit chains. Insertion of the cyclic shifts prevents unintentional beamforming. Windowing optionally smoothing the edges of each symbol, increase spectral decay. Carrier and sampling frequency offsets are deliberately simulated in the system. The data packets of all transmit streams are then passed through wireless fading channel followed by addition of Gaussian noise to it.

4.2.2 Modulation and Demodulataion

The FEC-coded bits are sent to the mapper that will map these bits to (BPSK, QPSK, 16-QAM or 64-QAM) constellation depending upon the requested rate, modulation and coding scheme. The Modulation and Coding Schemes (MCS) as specified by standard are listed in Table 4.2.2. On the receiver side, demodulation and FEC-decoding is performed after MIMO dectection/STBC decoding stage.

MCS Index	**Modulation**	**R**	$N_{BPSCS}(i_{ss})$	N_{SD}	N_{SP}	N_{CBPS}	N_{DBPS}
0	BPSK	1/2	1	52	4	52	26
1	QPSK	1/2	2	52	4	104	52
2	QPSK	3/4	3	52	4	104	78
3	16-QAM	1/2	4	52	4	208	104
4	16-QAM	3/4	4	52	4	208	156
5	64-QAM	2/3	6	52	4	312	208
6	64-QAM	3/4	6	52	4	312	234
7	64-QAM	5/6	6	52	4	312	260

Table 4.1: MCS parameters for mandatory 20 MHz, $N_{SS} = 1$, NES=1

4.2.3 Transmit Diversity

Employing multiple antennas at both transmit and receive side provides diversity gain and reduces the wireless channel effects by transmitting data simultaneously over multiple independent fading channels and then properly combining them in the receiver [109] [110]. As per standard, we use Alamouti encoding scheme at transmitter; and at the receiver we perform Alamouti decoding and receiver combining for space-time coded data. This produces two independent data streams that can then be transmitted simultaneously from two antennas over different fading channels.

4.3 Energy Consumtion Models

A number of works in the literature report measured statistics for WNIs and mobile phones. Although the actual numbers are dependent on the particular implementation, the power consumption trends over the configuration space mostly stay the same. We present the results reported in [1] and [107] in 4.3.1 and a comprehensive power consumption model in 4.3.2 which is based on the work presented in [2].

4.3.1 Measurements at Interface

In [1], power measurements for Intel Wi-Fi 5300 a/b/g/n Wireless Network Interface (WNI) was performed. This WNI supports upto 3 streams both at the uplink and the downlink. For our purposes we are only concerned with the transmit side of the interface, i.e. the uplink. The tests were conducted using two wireless nodes reasonably close to each other so that retransmissions are not necessitated and all rates are supported. The nodes were running 2.6.33-rc7 linux kernel and Intels's iwlagn driver [111] which was modified so that the configuration: RF chanins, channel widths, number of spatial streams etc could be controlled.

Instrumentation

The nodes in the experiment are available in the mini-PCI Express form factor which was connected to a PC through to min-PCI Express to PCI Express converter. One of the two nodes was instrumented A current sense resistor valued at 40mω, on the 3.3V power supply to the WNI. By tabulating the voltage drop across the resister, the current and hence the power consumption by the WNI was calculated. Packets of 1500 bytes were transmitted during the experiment.

Sleep vs Idle

In the idle 1-Rx mode, WNI is only listening for the transmissions on the idle channel. This is active mode with the lowest amount of activity. In the sleep mode however, the WNI only wakes up for the beacon signal or as a result of an interrupt generated by the host when it intends to communicate. Power consumption in both the mode are compared in table 4.3.1 Clearly, the power draw increases eight fold from sleep mode to idle mode.

Rate

At the transmit side, it was found that a change in rate had little effect on the transmit power requirements. staying in the same antenna configuration, only by changing the MCS does not change the transmit power requirements by a significant amount.

Mode	Power Draw
Sleep	100 mW
Idle 1-Rx	820 mW

Table 4.2: Sleep mode vs Idle Mode Power Draw

Mode	Power Draw
Tx SIMO	1280 mW
Tx MIMO 2-Tx	1990 mW
Tx MIMO 3-Tx	2100 mW

Table 4.3: Power consumption for various transmit antenna configurations [1]

Channel Width

Doubling the channel width, i.e increasing from 20 MHz to 40 MHz has little impact on the power consumption at the transmitter. This result implies whenever possible, wider channel widths should be used.

Number of Transmit Antennas

From the numbers in Table 4.3.1, it is evident that the increase in power requirements when going from 2 to 3 antennas is only a tiny fraction of the increase when going from 1 to 2. Transmitting with 2 antennas requires around 700 mW more power than transmitting with single antenna, and only a 100 mW lesser power over transmitting with transmit antennas. Therefore, when going beyond a single transmit stream, the maximum number of transmit antennas available should be utilized.

Transmit Power

Reducing the transmit power from 32mW to 1 mW would have a tremendously adverse impact on the received SNR. The decrease in the power consumption by the WNI however was only 140 mW. Similar results have been reported in other works as well e.g. in [104]. This suggests that the transmit power variation constitutes only a very small proportion of the overall power consumption for the WNI, thus the received signal degradation far outweighs the small gains in power consumption.

Normalized Power Consumption

As it was determined from the figures reported previously that power consumptions vary minimally for the same number of antennas when the MCS is varied, faster MCS has a lower energy costs when the power in Joules per second is normalized data rate in bits per second to determine energy per bit. Figure 4.1 depicts this comparison.

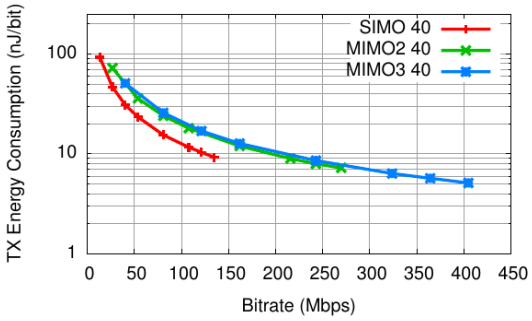


Figure 4.1: Normalized Power in Joules per bit [1]

Conclusion

These numbers supports the "race to sleep" strategy, i.e. the fastest configuration available should be used for transmission so that the transmitter can be in the sleep mode for as long as possible. For shorter packets, since the packet overheads for more antennas would form a bigger proportion of the packet length, the fastest configuration for a single antenna should be used. For longer packets, the maximum number of antennas should be used with the configuration supporting the fastest rate

4.3.2 Comprehensive Model

While the power measurements in the last section were performed on the interface, in [2] the power consumption is divided into its constituents and a comprehensive energy model is generated. It is reported that a substantial amount of energy is expended while processing the frame across the implementation stack. This energy, termed cross-factor energy, is found to be independent of the length of the data.

Although there is noteable power consumption when WNI is plugged in the slot, power consumption after loading the driver and associating with an access point is taken as the baseline ρ_{id}. Power consumption statistics are measured by varying frame size L in the range 100 to 1500 bytes, modulation and coding between 6, 12, 24 and 48 Mbps, transmit power between 6, 9, 12 and 15 dBm and frame generation rate λ_g upto 2000 frames per second.

Figure 4.2 shows the power/airtime plots for Soekris and Alix platforms. Total power consumed by the device is plotted against the percentage airtime τ_{tx} which is computes as

$$\tau = \lambda_g T_L \tag{4.1}$$

where λ_g is the frame generation rate and $T_L = T_{PLCP} + (H + L)/MCS$ is the time required to transmit a frame of length L, date MCS Mbps. T_{PLCP}

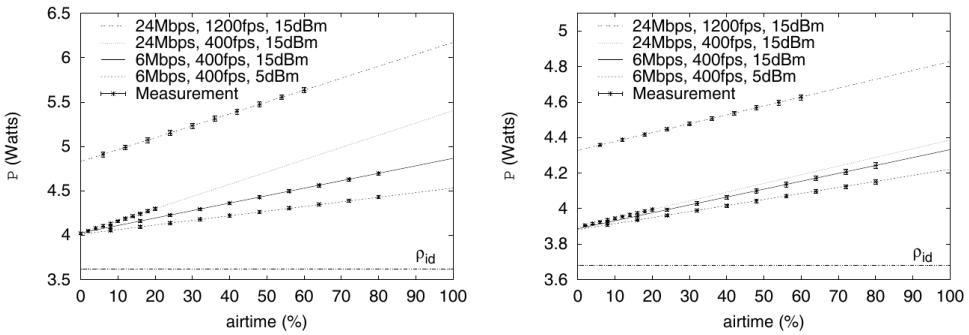


Figure 4.2: Total Power Consumped vs Airtime for a) Soekris Platform and b) Alix platform [2]

accounts for the transmit preamble and H for the MAC layer overhead. While the actual number for the two platforms in 4.2a and 4.2b are slightly different since it depends on the hardware, their variation over different operating parameters follows the same trend enabling the decomposition of total power draw into three constituents,

$$P = \rho_{id} + P_{tx} + P_{xg}(\lambda_g) \tag{4.2}$$

where ρ_{id} is the platform specific baseline power component, P_{tx} is the transmit power which grows linearly with airtime τ_{tx} and $P_{xg}(\lambda_g)$ is the large offset above ρ_{id} from the linear power consumption trend starts. This factor, which is a function of frame generation rate λ_g is the main contribution of the work [2].

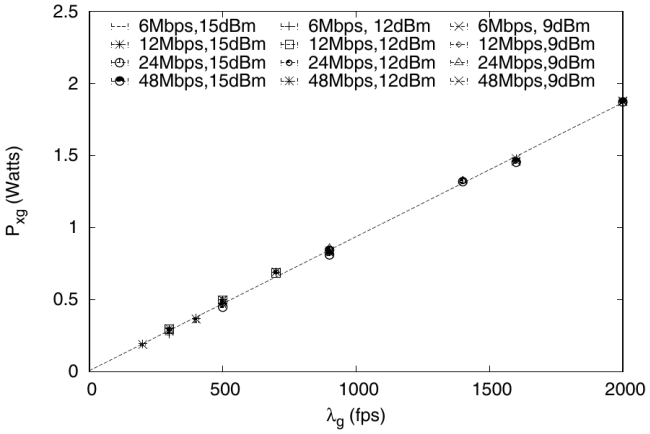


Figure 4.3: Relation between $P_{xg}(\lambda_g)$ and λ_g [2]

Plot 4.3 depicts the relationship between $P_{xg}(\lambda_g)$ and λ_g for various MCS and transmit powers. The plot clearly shows that $P_{xg}(\lambda_g)$ is independent of the frame

size and radio settings. The constant of proportionality $\gamma_{xg} = P_{xg}(\lambda_g/\lambda_g)$ is the energy toll for processing of each individual frame irrespective of the size and transmit power.

Cross Factor

By dropping packets at points a, b and c in figure 4.4, energy consumption up to the respective points was evaluated. Based on these evaluations, the plots 4.5 were obtained. It is visible from the plot that the energy toll due to frame processing is practically independent of the frame generation rate and frame size.

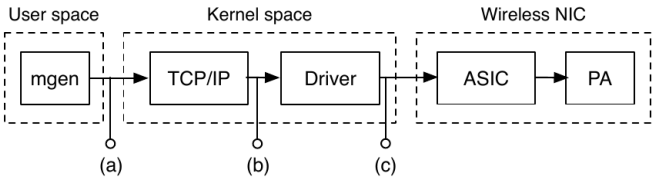


Figure 4.4: Interface Crossed During Transmission [2]

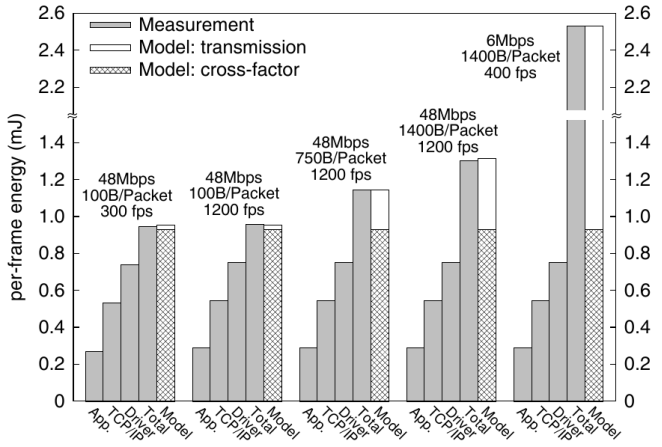


Figure 4.5: Per Frame Energy Cost in Transmission [2]

The per-frame processing cost appears to roughly split as follows: 24% application 33% TCP/IP stack, 21% driver, and 22% NIC.

The comprehensive power consumption model at the transmitter without retransmissions therefore is:

$$P = \rho_{id} + \rho_{tx}\tau_{tx} + \gamma_{xg}\lambda_g \tag{4.3}$$

where

ρ_{id} is the platform specific baseline power consumption $\rho_{tx}\tau_{tx} = P_{tx}$ is the transmit power, which is linearly dependent on the transmit airtime τ_{tx} γ_{xg} is the cross-factor while λ_g is the frame generation rate. The last term is the addition of this model to the traditional ones.

4.3.3 Effective Strategies

The effective strategy depends, when viewed for the wireless interface only, is highly dependent on the power consumption by the analog front end. Since the difference between consumed power in sleep, idle and active modes are quite significant the general strategy is race back to sleep. Power consumed by the baseband processor severely outnumbers that consumed the RF frontend. Additional power consumed by each added RF chain diminishes over every addition. Thus when using more RF antennas, generally using the maximum available antennas is more suitable. Attempting to control power consumption through controlling the baseband signal parameters is effective in only some of the cases. Specifically, transmit power control would have significant effects on BER rate performance but reducing transmit power has only marginal effect on the overall power consumption since transmit power constitutes only a tiny fraction of the total power consumption. The race to sleep strategy works in general but with some restriction: which is the number of RF antennas to be used. For shorter packets, highest MCS should be used with only a single RF chain. For large packets, the highest MCS should be used with the maximum RF chains possible.

When viewed for the whole device, another factor is revealed, namely the cross-factor. This the energy consumed above the device driver while the frames travel across the protocol stack. This factor turns out to be independent to the frame size, which is impactful when dealing with variable frame rates. Cross factor is dependent on the frame handling itself. Packet batching is an effective for minimizing cross factors. Packets can be batched together at the application layer and passed through all the layers before de-batching at the driver.

4.4 Strategies at the MAC Layer

The IoT paradigm is highly dependent on very strong computation platforms. Huge amounts of data gathered from billions of things is centrally processed at clouds. The uptake of cloud computing is already allowing the solution providers to reduce their energy consumption by optimizing their centralized computation resources. Transmission of data to these mammoth centralized computation resources is via wireless networks These wireless networks were not designed for energy conservation but rather for spectral efficiency maximization. With a rapid shift to cloud resources, the situation is changing so drastically that according to [112] by 2015 wireless networks will account for up to 90% of the wireless cloud computing energy consumption.

In IoTs, power consumption is not only significant from an environmental perspective, but also from an operational point of view as well. When the nodes in the IoT are multimedia devices or video cameras in particular, energy consumption in transmission becomes orders of magnitude more important. This is because the nodes, depending upon the deployment scenario, the nodes, being powered by batteries, may have to transmit continuous streams of visual data.

The evolution of batteries, unfortunately, has been far adrift of the Moore's law, currently only expected to attain only a two-fold increase in density over the next ten years [113]. Thus a great amount of research effort is already being dedicated for the developing power consumption models and for the creation of methods of reducing energy consumption [114] [115].

Lot of work have been done in the literature to reduce energy consumption of station while in idle state, such that the stations switch to doze state when they don't have any packet pending for transmission or reception. Few schemes which predict the idle time and scheduling beacons non-periodically to extend sleeping time are proposed in [116], [117], [118], [119], among others. In these technique authors argue that when stations are communicating through Internet, then the idle time and its intervals can be irregular and on-periodic. However, for real-time multimedia traffic these assumptions are not as such appropriate. In addition, these techniques considered only a single transmitting station, thus adapting dynamic beacon interval is hard in case of multiple stations. Similarly, in [120] the overhead of periodically waking up for the reception beacon frames is reduced. In the proposed scheme stations instead of receiving every beacon frame the stations adaptively select the number of beacon intervals after which they have to wake up. Depending upon the traffic level the intervals are modified like the Binary Exponential Backoff (BEB) selects contention window size i.e. if no traffic is scheduled for them in a given beacon they exponentially increase the wake up interval time. By comparing the scheme with PSM, the authors reported to save 42% energy in constant bit rate traffic. Modern hardware has enabled quick switching between on and off states of the radio, exploiting this in [121] authors proposed an aggressive power management scheme in which a station can sleep even for smaller durations. The packets lost during the short nap period are reported to be very small. However, the scheme is only applicable for downlink traffic scenario and thus not feasible for IoM.

Due to the enormous adoption of IEEE 802.11 based WLAN, usually the networks are dense. Similar level of density is expected for networks based on IoM. Higher number of nodes in the network results in increased interference and time spent in contention; as a result nodes get less time to sleep and wait longer for transmitting or receiving pending packets in queues. Dividing the beacon interval into time slots and then assigning these slots to individual nodes in a TDMA like mechanism is shown to reduce contention time and collision probability in [122], [123], among others. In this way, power saving nodes can sleep longer and wake up only on scheduled time slots. In [124], authors proposed another scheduling

mechanism named NAPman which control the traffic destined for PSM-enabled stations so that the other stations in the network don't starve. To enable energy efficiency as well as fairness among network nodes, the AP notifies traffic for PS-station only if there are no pending (older) packets for non PS-stations. This approach is proved to perform better in dense network scenarios when multiple stations communicate with a single AP.

The combined operation of frame aggregation and IEEE 802.11 power saving mechanism is promoted in [125]. The authors proposed a Congestion Aware-Delayed Frame Aggregation (CA-DFA) algorithm, in which the congestion level of the network is estimated and the transmission are intentionally delayed. When the congestion level drops to a certain threshold, then the pending packets are transmitted using frame aggregation technique. Consequently, the collisions are reduced and the energy is not wasted in retransmissions. However, this approach is not feasible for real-time multimedia traffic, since the scheme doesn't incorporate the user experience in terms of the delay induced by CA-DFA algorithm. In [126], an adaptive-buffer power save mechanism (AB-PSM) is proposed which tries to prolong sleeping time of the mobile stations streaming multimedia content. Authors proposed that an application buffer stores the pending packet and doesn't allow the AP to advertise them in the TIMs, so that stations can sleep longer. Although in this work the battery life of the mobile station is analyzed, however no QoS or user experience is considered while delaying the packets and also QoS experience performance is not evaluated. Another buffering and queuing technique named Low Energy Data-packet Aggregation Scheme (LEDAS) is proposed in [127]. Packet received from different applications are aggregated and transmitted. However, the authors in [127] didn't consider per application based delay bound or user experience.

4.5 Enhanced IEEE 802.11 MAC (IEEE 802.11+) for Low-Power Multimedia Devices

IoM implication in is true color requires resource constraints multimedia devices to be connected and accessible through the Internet. However, the multimedia content being communicates over the network requires different traffic requirements as compared to data communication such as bandwidth, delay, jitter and reliability. These network performance requirements are referred as the Quality of Service (QoS), which represents the level of user experience. For example, how fast is the data transmission, how much is the delay at the receiver, what is the probability of correct data reception at the receiver, what is the probability of the transmitted data to be lost, etc.

Most of the communication models and power saving techniques are designed considering the multimedia traffic to be downlink traffic i.e. end users retrieving multimedia content from or through the Internet. However, in many IoM scenar-

ios the end device or sensor is the resource constraint device which is responsible for generating multimedia content, i.e. video camera node, and transmitting or uploaded to the Internet. The IEEE 802.11 power saving mechanisms and most of the proposed schemes to enhance these power saving mechanisms do not consider this uplink multimedia transmission scenarios. The current standardization activities of providing Internet-connectivity to 'Things' are not focused to address the challenges of provisioning multimedia objects over Internet of Things. Many researchers have investigated a variety of techniques to limit the power consumption of wireless multimedia sensor networks. However, these issues have not been addressed considering WMSNs based on IoM architecture.

In our proposed communication methodology, we used a twofold technique to make IEEE 802.11 more power efficient. Firstly, we try to decrease the protocol overhead of the IEEE 802.11 to increase bandwidth efficiency which results in lowering the duty cycle and saving energy. Secondly, we try to schedule the multimedia data transmissions to decrease the duty cycle considering the user experience level requirements i.e. delay bound, video quality etc. To achieve these improvements, in our proposed power saving communication mechanism the stations adaptively select appropriate frame aggregation threshold and the TXOP duration based on its current data transmission rate and the QoS specified delay bound. In addition, the stations provide certain quality of the multimedia content depending upon its battery status. To prolong the life time of the network the low priority frames are dropped if the remaining battery of the station is lower than a specific threshold level.

The proposed algorithm is designed on top of the IEEE 802.11n PSMP power saving algorithm, by optimizing it for real-time multimedia traffic requirements. At a power saving station, when a data packet is received at the MAC layer from the higher layer, it is appended in the pending packets queue. If the current packet is the first packet in the queue, then its arrival time is retained. Depending upon the Channel State Information (CSI) received from the intended receiver node, the transmitter determines the highest data rate that can be supported in the given channel conditions. Higher data rate enables reduction in the time of reception and transmission of the packets, which in turn reduces energy consumption. In addition, higher data rate alleviates the need of longer PSMP-UTT duration requirement, thus bandwidth resources are efficiently utilized.

Prior studies on IEEE 802.11 access mechanism have already proven that the frame aggregation mechanism is a bandwidth efficient which decreases the protocol overhead, thereby reducing power consumption. However, the frame aggregation threshold i.e. the amount of data that can be transmitted in a single packet transmission is a critical network performance metric. Since, real-time multimedia transmission requires stringent QoS, in terms of delay, bandwidth, latency etc; therefore frame aggregation threshold should be selected or adapted according to the application requirements. Therefore, in our proposed scheme the frame aggregation threshold is estimated in accordance to the supported data rate and the

application specific delay bound with respect to the packet arrival time of the oldest pending packet in the queue. Once the amount of data that can be transmitted in an aggregated frame is known, the transmitter node waits for more packets till the data exceeds the frame aggregation threshold, till the delay bound for the oldest packet is acceptable. The pending packets in the queue are transmitted as soon as the amount of data in the queue exceeds frame aggregation or delay for oldest packet equals the delay bound.

In video compression techniques like H.264/MPEG-4, Scalable Video Coding (SVC) is employed, which comprises of the base layer and three enhancement layers i.e. temporal, spatial, SNR/Fidelity, these layers are used to scale the frame rate, resolution, quality of the video. However, if a particular base layer frame is lost then the relative enchantment layer frames are useless and therefore transmitting these frames will result in power wastage. Since, the video frames have different priority levels and are not mutually exclusive, thus for efficient bandwidth and energy utilization these priorities must be considered. For this reason, in our proposed scheme, we provide three levels of transmission priorities in terms of the Retry/Retransmission Counter value. Since, we are considering real-time video traffic, thus attempting retransmission of lost packets multiple times is of no use. Thereby, we provide different retry limits for I-frames (3), P-frames (2), and B-frames (1). If this limit is surpassed, then these frames are dropped and instead relatively newer frames are transmitted.

Once a transmitting node determined the aggregation threshold and achievable data rate, it can ask the PSMP-enabled AP for the TXOP or more accurately PSMP-UTT service period. In PSMP AP shares its own TXOP to provide PSMP enabled stations to transmit uplink traffic and/or receive downlink traffic. Since, the real-time video traffic application been considered here doesn't require bi-directional traffic communication, thus the PSMP-DTT is mostly used just for transmitting the BlockAcks to respective stations. By adopting the proposals given above, a station generating real-time video traffic can effectively transfer the multimedia content with the AP by efficiently utilizing service period allocated to it and stay in the doze state as much as possible, saving substantial amount of energy.

4.6 Network Configuration

In computer networks the network configuration refers to the arrangement of the network objects and distribution of the network resources to facilitate network nodes to communicate within and outside the network [128] [129] [130]. Firstly, network objects are arranged in a desired manner referred as Network Topology, such that the network is modeled as a graph with network nodes and the links between nodes being represented by vertices and edges, respectively. Once the topology is defined then the available network resources can be distributed to get the desired network performance optimization. The topologies can be physical

topology which depends upon the location of the objects and the logical topologies which depict the flow of the data in the network. In this work we only consider the logical topology for network configuration.

The network topology control is used in Wireless Sensor Networks (WSNs) and Mobile Ad-hoc Networks (MANETs) to reduce energy consumption by the individual node or the whole network, avoiding interference between network nodes, extending the lifetime of the network, etc [128, 130]. There are multiple basic topology models, i.e. point-to-point, ring, tree, star, mesh, bus, etc, which can be selected as per the desired data communication and application requirements. Network topology should be updated of modified in dynamic network environments to preserve network connectivity, density, coverage, etc. To control the network topology and manage network operation a network administrator can typically control the following parameters; network resources like bandwidth, transmit power of the nodes, states of the nodes i.e. active or sleep, roles of the nodes i.e. gateway, cluster-head, regular.

ZigBee over IEEE 802.15.4 standard [33] defines specifications for low-power devices in a Wireless Personal Area Network (WPAN). To extend the lifetime of the network, the energy sources of the nodes are efficiently utilized i.e. duty cycle for these low-power devices is minimized (typically less than 1%). Two different types of network topologies are specified, star topology and mesh topology, enabling multi-hop, reliable and self-organized communication. In addition, three different roles for network devices are specified i.e. network coordination, regular operation and relaying, or simple regular operation. For these roles two types of devices are defined Full-Function Device (FFD) and Reduced-Function Device (RFD). FFD can serve any role i.e. as WPAN coordinator, relay node, or regular device. However, there can be only a single WPAN coordinator, FFD device can be multiple depending upon the network density and application requirements. On the other hand, the RFD device is much simpler and it can operate only as a regular device performing simple operation. FFD devices can talk to any other device in the WPAN, however, the RFD devices can only talk to WPAN coordinator or FFD device (can be associated with only one at a time) and not with other RFD devices in the network.

RFD devices have small memory sizes and simple computational capabilities, since they are designed for simple applications like operating the light switch or sensing scalar data. Thereby, the RFD devices usually have only small amount of data which can be swiftly transmitted or reported to the FFD or WPAN coordinator, and the device can switch back to sleep mode to save energy resources. However, in mesh network topology the FFD devices have to stay awake much longer to provide connectivity between RFD device and WPAN coordinator as compared to the star network topology. To establish a WPAN, initially a FFD device may choose itself to become the WPAN coordinator. After that it can broadcast the existence of a WPAN network coordinator in its vicinity. The other devices hearing availability of the WPAN, associated with the network by sending

joining requests to WPAN network coordinator. Once associated with the network, these devices can operate as per their designated roles. In addition, to WPAN coordinator other FFD devices can also provide coordination and synchronization services to their respective connected RFD devices.

Once the network topology is established, now we a routing protocol to enable data communication. IoT incorporates battery operated low-power wireless device which experience lossy radio links due to multi-path fading, topology changes due to mobility. For this reason, the IETF working group Routing Over Low power and Lossy (ROLL) networks adopted a gradient-based approach [38], [54], [47], [55], to design an effective IPv6 enabled Routing Protocol for Lossy Link Networks (RPL). This routing protocol RPL is compatible with multiple link layer technologies, specifically low-power based link layers i.e. IEEE 802.15.4. Thus, it is suitable applications in building/home automation, industrial environments, and for other applications [56, 131, 132, 57].

In a RPL based network scenario, some device is opted to be root device and collect data by coordinating with other network devices through multi-hop routes. Root device creates a Destination Oriented Directed Acyclic Graph (DODAG) for each other device based on the link costs and other device attributes. Then using an identifier DODAGID, this objective function is optimized for required conditions. Also, based on the distance between a device and its respective root device a Rank metric is build which is used to set up a topology. The Rank monotonically decreases with the distance as per gradient-based approach i.e. farther the device from the root lowers the Rank and vice versa. The RPL protocol determines both the downward routes from root or gateway devices to end sensor or actuator devices for Point-to-Multipoint (P2MP) as well as the upward routes for the other way around which is used for Multipoint-to-Point (MP2P), these routes may also be used for Point-to-Point (P2P) communications as well.

The RPL topology can be as simple as network with only a single root device maintain a single DODAG, though this kind of implementation is feasible only for small number of network device. However, in case of a large network root device may maintain multiple virtual root devices responsible for only a part of the network. In this way, the DODAG becomes simple and flexible and the maintenance of the network can be done in efficient way. A more sophisticated and flexible configuration could contain a single DODAG with a virtual root that coordinates several LLN root nodes.

RPL routing protocol can exploit key network metrics such as: node energy, hop count, link throughput, latency, link reliability, and link color, to effectively adapt to the network channel conditions [58]. Specifically, RPL uses a metric (link color) to estimate based on the color of the link whether to include or exclude it for a particular DODAG. These metrics can be shared with other devices and the routes can be adaptively selected based on network conditions. However, more frequent changes results in instability in routes and degrade network performance.

Generally, while deploying IEEE 802.11 based WLANs the capacity enhance-

ment is given much consideration as compared to energy efficiency. However, to devise IEEE 802.11 energy efficiency comparable to IEEE 802.15.4, so that it can be adopted for networks in Internet of Multimedia (IoM). Usually, all the nodes in the network operate with full transmitting power and stay continuously awake. We propose that for an infrastructure based IEEE 802.11 WLAN, not every node communicates with the Access Point (PC) directly. Instead they can talk through intermediate nodes towards the AP. In this way, far away nodes (with respect to the AP) in the network don't have operate in full transmit power. In addition to the energy saving, the decreasing transmit power comes with reduction in interference among network stations and this reduction increases the possibility of multiple links being active at the same time.

To enable energy efficient real-time multimedia over IEEE 802.11, we propose that the devices in the network i.e. End-Node, Relay-Node, AP, have different roles to increase the energy efficiency of the network. End-Node is responsible to acquire multimedia content from physical environment and transmit it to its respective Relay-Node or directly to AP (if AP is its vicinity), the Relay-Node can also acquire multimedia content and it is responsible to forward the multimedia content received from End-Nodes to the AP, the AP receive data from network nodes and report to a server or end-user via Internet. We consider that the AP is continuously provided with a power supply, so power is not an issue for AP, which is also usually the case. However, the End-Nodes and Relay-Nodes are equipped with multiple and heterogeneous energy sources i.e. from wind, thermal, vibrations, fuel cells, solar cells, AAA batteries. In the vicinity of an End-Node, if multiple nodes are capable of operating as a Relay-Node then only the one with most Green energy source will be selected i.e. the one with solar energy. Consequently, the overall CO_2 emissions by the network will be minimized.

Employing this network configuration strategy will enable network scalability, since low-power devices farther from the AP can communicate through intermediate Relay-Nodes. Similarly, our proposed tree topology will enable point-to-point communication between devices, simplifying routing management and nodes integration. In addition, the hierarchy in the topology makes network resources distribution easier i.e. a Relay-Node forwarding for larger End-Nodes can be allocated higher network resources in terms of banwidth or transmission opportunity and vice versa.

Chapter 5

Video Coding

A scene of a real-world is spatially and temporarily sampled in a digital video. The video is sampled in time which generates frames, that represent a scene at a point of time. Samples altogether produce a video signals. To represent a video signal, three components are used, which are:

1. High-Definition formats
2. Intermediate formats
3. ITU-R 601 standards

Digital video can be represented in form of RGB and YUV/Y4M [133]. Three colors red, blue and green is used to present RGB format. Color intensity (chrominance) and brightness (luminance) can be also adjusted separately in the RGB format. These can be added through calculating weighted sum of the colors thoses are R, G and B. YUV have handover on RGB format, as it can be represented with lower resolution, and not harmful to human visual system (HVS) caused by intense brightness. RGB format can be converted into YUV through following formulas

$$Y = (0.257 * R) + (0.504 * G) + (0.098 * B) + 16$$
$$U = (0.439 * R) + (0.368 * G) - (0.071 * B) + 128$$
$$V = -(0.148 * R) - (0.291 * G) + (0.439 * B) + 128$$

A source image or video sequence is converted into compressed form by video CODEC encoder and a copy is generated at the decoder end. The whole process is considered lossless, if the video received after decoder is identical to the original.

5.1 Video Processing

There is no specific technique to encode digital video described in the standard. Each video frame/sample is in spatial-temporal coordinate. The converted video samples after transformation is in form of bitstreams. Usually the bitstreams include certain information that is useful for decoding after prediction. Also defines compressed data structure and tools (profiles) used in the encoding phase including video sequence information. Following variables are used.
1) Syntax Elements
2) Variable Length Coding
3) Arithmetic Coding

Typically a video frame is split into two fields, those are temporally and spatially encoded. For encoded color images, YCbCr spaces are used separating into luminance and chrominance planes. Basically video encoding has three main operations listed below.

5.1.1 Prediction Model

For reducing redundancy, the aim of prediction is to subtract the prediction from fresh data. Residual samples is the output after prediction phase, energy or power consumed is less after prediction [134]. The samples are further processed in encoding phase. There can be of two types of predictions

• **Temporal Prediction**
In this type of prediction, the previous frame is the predictor of the current frame. Difference or residual samples are the output after the accurate process of prediction. Usually in residual less energy is contained. In temporal prediction past or future frames are compared to create a new frame known as reference frames. The residual sample is the difference of frame 1 and frame 2. Although in this type of prediction, alot of energy is consumed, usually due to moving objects of two frames.

• **Spatial/Intra Prediction**
In this type of prediction image samples are created in the block from previously-coded samples of the same frame. Different types of scans are used for reading data according to the situation, as shown in the below figure.

5.1.2 Image Modeling

Grid of sample values combined to form a natural video and images [135]. Because of the high correlation values between neighboring image frames. For a perfect motion compensation / intra prediction, the correlation values are less in residual image, as shown in below figure.

The properties of image modeling section are to decorrelate image and through entropy coder convert it into compatible compressed form. Following are the subsections of image modeling.

a) Predictive Image Modeling
b) Transform Modeling
c) Quantization
d) Zero encoding and Reordering
Autocorrelation is performed between neighboring samples, to attain height is graph, where similarity exists between shifted copy and original image. The height in the graph shows zero shift. If a gradual slope is attained meaning video samples exists in the neighborhood.

5.1.3 Entropy Coder

Series or varied symbols are converted into bitstream for transmission purpose by going through entropy encoder. The symbols may be of following types

i) Run level
ii) Zero tree structure
iii)Quantized transformed co-efficient
iv) Motion Vectors
v) Pixels values
vi) Color or marker codes
vii) Macroblocks headers
viii) Sequence / picture headers

Entropy encoding can be completed either through predictive coding or by variable-length coding. In the former approach in case of h.264 macroblocks frames use offset of reference prediction with previously encoded frames. Both are correlated with neighboring blocks extending till large regions. A approach Motion Vector Difference (MVD) calculates difference between actual motion vector and predicted one when its encoded and transmitted. If variable-length coding is used codewords are compared with the input symbols, codewords length may vary and compression is applied if the redundant symbols occur.

5.1.4 Transform Coding

Compressive Sensing of Image Reconstruction vs DCT

Classic sparse transformation includes discrete cosine transform (DCT), discrete Fourier transformation (DFT), discrete sine transformation (DST) and discrete wavelet transformation (DWT), the wavelet transformation [136] is used to change

the image to sparse signals. DCT related compressive sensing procedure changes image pixels in frequency domain [137], centralizing energy to lower frequency.

Algorithm Steps:
1) Transforming image pixels into 2D frequency domain matrix viva DCT transform, classifying those four sub-band LH1, HL1, HH1, LL1 very similar to wavelet decomposition.
2) Selecting a value V to constructing VxN/2 random size matrix having values distributed upon Gaussian probability density function with variance 1/N.LH1, HL1 and HH1 and mean zero. Obtaining three sub-band measurement values.
3) The frequency coefficient matrices are reconstructed viva 2D IDCT transform for following matrices LH1, HL1, HH1 and LL1. Shown in below figure.
Another approach for reconstruction of image block sparsity is used for CS to enhance the detecting sample plain. Using block

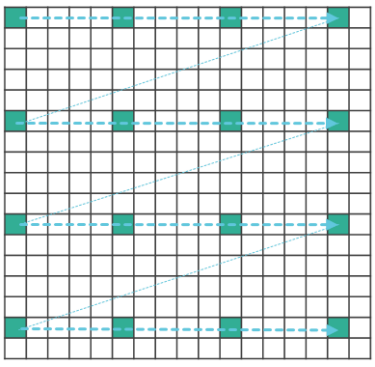
$$G(x) = \sum_{i} |X_{Bi}| \tag{5.1}$$

Where (Bi)i is a segment of disjoint indexes 1,.,N, where extracts the coefficients within B and is the norm. The proximal operator of this block norm is a block thresholding. Following are the complexity cost of usual approaches for handling image reconstruction.

5.1.5 Scan Techniques

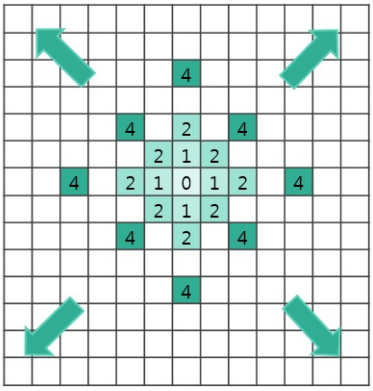
For motion compensation or estimation, when encoder starts to read a certain frame block, there are number of different scans [138], three of fast hierarchical approaches present listed below

• Rastor Search or Scan
In this scan technique, first row of the block is scanned from left to right, later again from left to right in second row and this pattern continues. In this search technique data re-usability ratio increases relatively but with redundant loading.



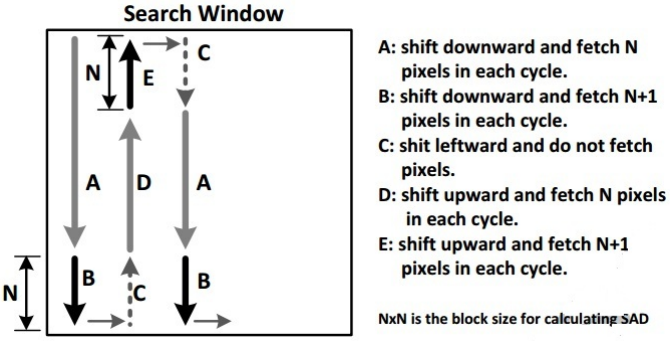
(5.2)

• Diamond Search or Scan
For multiview video is a three dimensional approach, maybe captured through number of different cameras. For reducing computational complexity multi grid diamond search pattern can be deployed according to preference as shown in following figure.



(5.3)

• Snake Search or Scan



(5.4)

5.2 Low-Complexity Algorithms

The objectives of low and scalable video encoder are to efficiently split-up a single video stream into multiple packets usually called layers or frames. At the receiver, the receiver has the option to select arbitrary layers to decode and process, as they arrived at the receiver end. Previously several scalable video codec standards [139], such as MPEG-2, H.263 and MPEG-4 were introduced. Although these mentioned standards did not provide coding efficiency, if compared with certain non-scalable video profiles such as Schwarz/Oelbaum, 2007. Further on, following are the selected current low complexity scalable video encoding approaches some are discussed in this section.

5.2.1 Low Complexity Encoder Based on GOP Methodology

The prediction residue is put through a spatial 2-D transform and the variances of the transform components are estimated using the coefficients. The bit-allocated are used to design scalar quantizers that operate independently on the transform components.

Global Motion Model

The global motion consists mostly of scaling, rotation and translation [140]. Mathematically it can be expressed as

$$T_\theta(x,y) = \begin{bmatrix} \theta_3 & -\theta_4 \\ \theta_4 & \theta_3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$$

Where (x,y) in frame are position of the pixels, the parameters for describing the rotation and scaling are

$$[\theta_3, \theta_4]^T / epsilon R_2, T_\theta(x,y)$$

These the position of the pixel in the motion compensated frame, R is the translation vector. Spectral entropy is performed by bit allocation. M x N are total coefficients and only L are coded. The energy of the component λ_i is directly proportional to the number of coefficients η_i coded.

$$\eta_i = \lambda_i \sigma_2 * L \tag{5.5}$$

Average number of bits spent to code i component, is the number of bits required to code the binary significance map related to significant coefficient. The distortion in coding is caused by two sources
1) Quantization
2) Discarding coefficients
In the nutshell based on the relative energy GOP and the target average distortion, total numbers of coefficients are coded in frame. In the end, finally bits are allocated to them to meet the target distortion and coefficients with largest magnitudes are chosen. Mathematically expressed as

$$d_i^s = n_i * \text{E(quantization error)} + (N - n_i) * \text{E(energy of discarded coefficient ts)}$$

$$d_i^s = n_i * h_i / lambda_i 2^{-2bi^s} + (N - n_i) * lambda_i$$

5.2.2 Fast Mode Decision Encoding in H.265

H.265 standard is block-based hybrid video coding framework. Basically there exists three units,

1) Coding Unit (CU)
2) Prediction Unit (PU)
3) Transform unit (TU).
This technique [141] of structuring is helpful for optimizing. In this method intra prediction is implemented to reduce the effective spatial redundancies with one image. By referring to the neighboring samples from all 3D dimensions that is left, up and top-right region. Theoretically it can provide 34 prediction modes [2]. Mathematically estimating the probabilities of this algorithm can be defined as

$$P(M_{curr}|(M_A, M_B)) = P(Mode_{curr} = min(M_A, M_B)|Mode_A = M_A, Mode_B = M_B) \tag{5.6}$$

Where $Mode_{curr}$, $Mode_A$, $Mode_B$ are random variables that represent the RD optimal prediction mode. The intra prediction directions when accessed in angular predicted mode have angles of +/-[0,2,5,9,13,17,21,26,32]/32.
In this technique rough mode decision is combined along with RDO for detecting best intra direction. Number of direction is intra prediction is reduced and further correlation between the neighboring blocks comparing with target block is used to clear the loss of coding efficiency at first phase. Five levels are specified related to number of N modes through which best are to be selected.

5.2.3 Optimized Distributed Video coding encoding via Turbo Codes

Regular key frames, K are subset of frames which are encoded and decoded based on 8x8 DCT codec [142]. For frame S, each pixel value is uniformly quantized with 2M intervals. Using adjacent reconstructed frames, the side information is generated by motion-compensated using Ziv frames by interpolation. WZ codec [143] is used to decode and encode only the even frames. Using a uniform quantizer with each pixel u of F2i is quantized using a uniform quantizer with 2M levels to produce a quantized symbol **q**.

As a rule, a video signal has high temporal redundancies due to the high correlation between successive frames. Using 2D transforms, exploring the temporal redundancy of the video at certain approaches such as 3D (Three Dimension), and recompense step is computationally avoided. By this approach the transformation converts the temporal and spatial correlation into high spatial correlation. A high video compression ratio is achieved by using de-correlation by resulting pictures DWT (Discrete Wavelet Transform) generating efficient energy compaction.

5.2.4 Low Complexity H.265 Encoding based on Level and Mode Filtering

In H.265 there are about 34 intra prediction modes. And according to prediction unit size [144] , the support mode varies. Following are the main steps

Main Algorithmic steps
1) Angular directions are set according to the candidate mode with the minimum HAD (Hadamard Transform Absolute) costs.
2) The best prediction is detected from all the candidates if it have minimum mode R-D (Rate-Distortion) cost, mathematically expressed as

$$JRDO = SSD + .R \tag{5.7}$$

Where R is denoted as total bits for encoding process, is quantization parameter (QP) and SSD is sum of squared difference.
Fast Pre-processing Stage
Complexity is involved for R-D intra mode-decision, resultantly great computation is required making it an unsuitable technique. Taking a simpler HAD-based cost function

$$JHAD = SATD + .R \tag{5.8}$$

Where SATD is represented as the sum of absolute transformed differences providing an accurate cost evaluation. Complexity is reduced by eliminating, data dependency likely to ease the design of parallel and hardware software. [141] Usually each CVSATD has 34 elements and each element estimated under the corresponding intra mode (from 0 to 33) original pixels, through which the SATD cost is calculated. Firstly, the CVSATD is based on 8x8 units for larger PUs rather than calculating them straightforwardly. Mathematically the estimation can be expressed as

$$\begin{aligned} CV_{SATD64} &= \sum_{i=0}^{63} CV_{SATD8\_i} \\ CV_{SATD32\_j} &= \sum_{i=0}^{15} CV_{SATD8\_i+16*j(j=0,1,2,3)} \\ CV_{SATD16\_j} &= \sum_{i=0}^{3} CV_{SATD8\_i+4*j(j=0,1,2,.....15)} \end{aligned} \tag{5.9}$$

Simplifying the above equation, it can be rewritten as

$$JHAD = SATD' + \Delta * E + \lambda * R \tag{5.10}$$

Where $\Delta * E$ is the estimation error and SATD‘ is the estimation to SATD. SATD‘ is less in quantity as compared to SATD because it is generated from prediction based of small blocks using reference pixels whom are likely to be more similar to the current pixel. Thereby generating $\Delta * E$ positive. Usually increases with having a size of 2N of prediction unit, since for the lower 8x8 level to higher 2Nx2N levels for the estimation of SATD‘. Four NxN units are divided through 2Nx2N unit recursively by cost of adjacent levels. To minimize the rate distortion, a set

of thresholds are defined whether on one 2Nx2N unit is split or not. Categorizing threshold for error rate is given below

$$E(T) = \int_{\Delta C_{2N}=0}^{T} P(split \mid \Delta C_{2N}) + \int_{\Delta C_{2N}=T}^{T} P(non-split \mid \Delta C_{2N}) \tag{5.11}$$

Where T is the threshold and E denotes error rate. Best threshold when $dE/dT = 0 and d^2E/dT^2$.

5.2.5 Low Complexity background modeling

In this approach background is used as the prediction reference, as this technique is very beneficial for achieving high efficiency [145]. Low complexity is attained by introducing a segment-and-weight based average method. Suppose there is an ideal background C‘, moving object M_{obj}, system noise N_{sys}, and a moving background M_{bgd} in a scene, by formulating the observed values from a sample scene V_{obsv} by:
$V_{obsv} = C‘ + N_{sys} + M_{obj} + M_{bgd}$
The symbol + is represented by cumulative effect. Ideal background $C_i‘$ for the i-th frame is derived by

$C‘(i) = V_{obsv}(i)N_{sys}(i) - M_{obj}(i) - M_{bgd}(i)$

The background frame suitable for object detection should be constantly updated utilizing the original input frame. In surveillance video coding system, each background frame in encoded into stream to guarantee the decoding match. The Bg frame in surveillance video coding for the n frames for detecting best, it should satisfy LGOP

$$\underline{Bg} = \arg_b \min \left\{ \sum_{i=0}^{n} J(b, V_i^{obsv}) \mid b \in B \right\} \tag{5.12}$$

Where in background frame, b is arbitrary in the available set of background frames B, and J(b,Vi) is a function for calculating the rate-distortion cost of encoding the i-th observed scene Vi with b as reference. This is called periodically background updating. The basic step is to calculate the threshold for segmenting. According to the distribution theory that is

$$f(x) = \frac{1}{\sqrt{2\lambda\partial}} e^{\frac{(x-\mu)^2}{2\sigma^2}} \tag{5.13}$$

Where σ is the mean square error, μ is the mean value.

5.3 H.264

As discussed above the purpose of any simple video encoder is to create a multi-stream components from a single-stream video. The multi-components are referred as layers. The H.264 standard was published in 2003. H.264 is based on MPEG-4. There are number of advantages provided by H.264 video coding standard [135]. H.264 provide support for bit rate adaptation that increases flexibility for packet level transmission at the NAL (Network Application Layer). Although it is video compression method for converting real time digital video sequence into bitstreams that requires less transmission capacity. By standardizing any specific technique, the manufactures of different products share the same platform and become compatible. H.264 was introduced and developed by a partnership between International Telecommunication Union (ITU-T) and International Organization for Standardization/International Electro-technical Commission (ISO/IEC). The H.264 uses pre-defined tools available in decoding phase termed as profiles. Each profile of H.264 defines certain tools. In encoder phase also H.264 have variety of tools available. For transmission it provides number of support for illuminating transmission errors and packeting in compressed format.

5.3.1 Coding Fragments Of H.264

H.264 profile have following coding parts in general.

1) Intra-coded slice (I slice).
2) Predictive-coded slice (P slice).
3) Context-based Adaptive Variable Length Coding (CAVLC) for entropy coding

The architecture of H.264 increases the capabilities offering solutions for encoding phase that support temporal, spatial and SNR qualities.

Encoder Complexity

The complexity for H.264 encoder is more complex rather than any previous video coding standard. There are computation and memory overheads in the development of embedded encoder/decoder. It does not offer any compatibility with the older version or video coding standards. Traditional approach is best defined in sensor transmission stream. Generally there exists an asymmetric complexity in the encoder and decoder phase in the video coding scheme [**?**]. In many applications such as multi-camera scenario and multi-user conference, with low processing capabilities, the complex compression has to be done in the processor. Usually distributed video coding has the base of punctured turbo codes. For the video coding

the turbo codes uses three approaches

1. Parallel concentration
2. For better weight distribution, interleaving
3. To maximize the gain and to enhance soft decoding decisions for decoding procedures.

To generate scalable compressed video bit streams, punctured turbo codes are applied. To get overall performance improvement, usually following steps are performed

• Predict a new frame from a previous frame and only specify the prediction error.
• Prediction error will be coded using an image coding method.
• Prediction errors have smaller energy than the original pixel values and can be coded with fewer bits. In this standard the system is broken into two layers [146] : namely

1) Network Abstraction Layer (NAL)
2) Video Coding Layer (VCL)

Video information is transformed into bitstreams through VCL, after the conversion the NAL layer maps the transformed bitstreams into NAL units those are HDLC-like and byte-oriented transportation-layer before delivery. Usually the NAL header is composed of three fields. NAL unit header is the first byte after the NAL unit code prefix. The bits are categorized as following.

1) Forbidden-bit (1 bit)
2) NAL-storage-idc (2 bit)
3) NAL-unit-type (5 byte)

forbidden-bit is used to indicate whether the NAL unit is corrupted or not. *NAL-storage-idc* is of two bits, which have relative importance, when the picture is buffered. The *NAL-unit-type* is of one-ten bits according to need. Usually they are referred as reserved bits. Particularly three operations are performed during the mapping phase those are

1) Byte Alignment
NAL unit adds byte header, that defines the categorization of the NAL unit. Synchronization is performed in the NAL unit through byte sequence.
2) Emulation Prevention
Emulation prevention bytes are used to interleave NAL unit payload data [147]. Those bytes prevent accidental data generation within the payload. These bytes

are termed as *start code prefix*, which are inserted in the data pattern. NAL unit structure are used in both bitstream-oriented and packet-oriented transport systems.

3) Framing

H.264 slices frames according to five types, which are listed below

• I(Intra) Slice: This type reference only itself. Usually it is the first received image of the video sequence or a still image. All the first frames from video sequences needed to be built from I-slices

• P(Predictive) slice: It takes reference from decoded or predicted slices for construction of video sample/image. Prediction is mostly not accurate therefore some residual images may be added.

• B(Bi-Directional) Slice: It takes reference from future and former P or I slices, except of that Bi slices are very similar to P slices. For the reason, I an dP slices are decoded after B slices.

• SI and SP (Switching) Slices:Usually they are used for transition between video samples of different natures and type. There probability or usage is although very less.

5.3.2 H.264 Main Modules

Spatial dependency layer in H.264 have the essential requirement of its own prediction module for performing intra and motion-compensated predictions within the layer. Another module that manages the scalability of quality is termed as SNR refinement of module. In the inter-layer prediction module, the dependency is managed between spatial layers, by reusing residual signal and motion vectors so to improve compression efficiency. All the modules are finally merged in a multiplex, a single integrated scalable bitstream having different spatial, temporal and SNR levels are defined. The process is combination of forward and inverse (encoding and decoding) path as shown in below figures the video frames are structured in macroblocks upon which perdiction is performed using either Inter or Intra predictions. After the transformation and quantization process, the resultant product is forwarded to **Entropy Encoder Module**. The output packets are finally formed in the **Network Abstraction Layer (NAL)** module. Decoding or Inverse path involves reconstruct of Macroblock data from previously transformed module which consist of **Deblocking Filter** [148] and **Transform and Quantization(ITQ)**.

Encoding Modes

Generally there are six encoding modes presented in academic and commercial bases, which are listed below

1) Single Pass-Constant Quantizer

2) Single Pass-Constant Rate Factor

3) Single Pass-Bitrate

4) Two Pass- Average Bitrate
5) Two Pass-File Size
6) Lossless Mode

5.3.3 Prediction Model

This section individually explains the characterization of all types of scalabilities [149]. It is generally suitable if the encoding and decoding complexity tallys itself proportionally with accordance to temporal and spatial resolution. A hierarchy should be established for video compression tools which involves scalability complexity.

Temporal Scalability

In this category video stream is transferred as subset of bitstream. Video frames are classified as three distinct types: I (intra), P (predictive) and B (Bi-predictive) as already known. In this procedure, three types of simplified motion compensation, termed as Temporal Prediction.

Temporal Prediction

• **No Motion Compensation**
Work well in stationary regions

$$f(t, m, n) = f(t - 1, m, n) \tag{5.14}$$

where t=time period, m=macroblocks and n=number of frames / block
• **Uni-directional Motion Compensation**
For uncovered ranges, its performance does not work well, as new objects are appearing

$$f(t, m, n) = f(t - 1, m - d_x, n - d_y) \tag{5.15}$$

where t=time period, m=macroblocks and n=number of frames / block, x and y are the coordinates
• **Bi-directional Motion Compensation**

$$f(t, m, n) = w_b f(t - 1, m - d_{b,x}, n - d_{b,y}) + w_f f(t + 1, m - d_{f,x}, n - d_{f,y}) \tag{5.16}$$

Spatial Scalability

Layered structure is used in spatial scalability; by layering the improvement of lower layer resolution is achieved. In this domain previously there have been two

prediction types introduced. 1) Extended Spatial Scalability
2) Inter-Layer Predictions
a. Inter-Layer Motion Predictions
b. Inter-Layer Intra Texture Predictions
c. Inter-Layer Residual Predictions

5.4 H.265

H.265 i-e H.265 is very much similar to H.264 AVC standard but alot have made easier. H.265 was a research project by Joint Collaborative Team (JCT-VC) working with joint effort between ISO/IEC and ITU-T. This literature will discuss the standard of H.265.

5.4.1 Levels Of H.265

There are three profiles related to H.265 for development of certain varied applications. Following are the profiles.

1) Main
2) Main 10
3) Main Still Picture
By limiting the number of profiles [150], interoperability is increased between devices. 4:2:0 chroma sampling is provided in these profiles, and does not use wavefront parallel processing. In these profiles 8-10 (B per sample) are provided. In Main Still picture only one coded picture is processed along the entire bitstream.

5.4.2 Design Aspect

Like H.264 H.265 is not structured as macroblocks, they are divided into coding tree blocks abbreviated as CTBs. According to stream parameters that can in depths of 64x64, 32x32 or 16x16 are hierarchy is arranged in tree structure having certain nodes. H.265 is categorized of having hybrid coding architecture. Basically tree blocks are explained in a way that an array is defined for a certain block having varied samples and sizes, encapsulating corresponding chroma blocks and luma together with syntax to code these. CTBs are split accordingly in a loop fashion in a Quad-Tree Structure. The limit till which it can be broken down is 8x8. For example a CTB of size 32x32 can be partitioned into four 8x8 regions and three 16x16 regions. Basically in H.265 three block structures are defined those are

1) Coding Unit (CU)
2) Prediction Unit (PU)

3) Transform Unit (TU)

5.4.3 Coding Tree Blocks

Coding Tree Blocks (CTBs) are included instead of macroblock in coding tree unit (CTUs). For the purpose of prediction [151], each CU included more entities creating transform (TU) and prediction unit (PU). The varied adaptive approaches are suitable for big pixel size, such as 2k x 2k in H.265 applications. A picture is divided into CTUs (Coding Tree Unit) in H.265 standard in encoder phase. CTUs are the replacement of macroblocks. The height and width of CTU are provided in a sequence parameter set such as 64 x 64, 32 x 32 and 16 x 16. Coding Tree Unit is defined as logical unit, consisting of three blocks those are two chroma samples (Cb and Cr) and luma (Y), the rest are associated syntax elements as shown in below figure. CTB s are contained in the CTUs. All the CTBs have same sizes as of CTU those are 64 x 64, 32 x 32 and 16 x 16. Inter/Intra prediction is decided accordingly with the CTB size. They are split into multiple sizes of coding blocks (CB) and predictions are decided upon them accordingly.

5.4.4 Advantages Of Quad Tree Structure

Macroblock is concept of basic processing unit for the video coding and termed as *coding tree block (CTB)*. A quad tree structure is a particular type of data structure, having internal node of four children. There characteristic are to partition two-dimensional space recursively into four regions. The regions can be of rectangular, square or any other shape. There are indexes attached with quadtree structure indicating how the block can be further subdivided for residual and predictive coding. In such a way adaptation can be easily done with the characteristics of signals. In decoding phase of quadtree structures the complexity lessen, as tree can be easily traversed viva depth-first approach by applying diamond or raster scan. Merge mode is introduced by h.265, adjusting all motion parameters in a inter frame of predicted block proportional to the merge candidate parameters. Through this structure, CABAC is implied only in entropy coding phase, not CAVLC. Wavefront parallelization tools can be used for efficient encoding, including improvement for deblocking filter and creating second filter named Sample Adaptive Offset [152]. A better efficient way to clear inter-frame redundancies Adaptive Motion Vector Prediction is used.

Slicing

In H.265, the frames are sliced very much similar to H.264. During scanning slices are represented as a group of *Largest Coding Unit (LCU)*. Slicing possibly can be used both for parallel and network packetization. When slicing is used although

rate distortion does occur. Because of rate distortion, new parallel processing techniques have been recently proposed.

5.4.5 Prediction In H.265

• Intra-Prediction

Prediction is done by neighboring blocks samples that are reconstructed. Modes are the same as of H.264 those are plane, DC, vertical/horizontal and directional [153]. In addition to those modes H.265 introduces row and column to maintain continuity. Angular mode prediction is complex in h.265 as compared to h.264 as multiplication is required. Mathematically expressed as

$$((32 - w) * x_i + w * x_{i+1} + 16) >> 5$$

Where w is a weighting factor and is reference sample. The weighting factor when predicted row or column wise remain constant accordingly to (SIMD) Single-Instruction Multiple-Data implementation.

• Inter-Angular Prediction

In AVC spatial-domain intra prediction successfully been used. Due to the increased size of the tree blocks, the intra prediction of HEVC operates in the spatial domain, thereby increasing in selected directions [26]. HEVC supports a total of 33 prediction directions named as $Intra_Angular[k]$, where mode number is denoted by k ranging from 2 to 34. On purpose the angles at near-vertical and near-horizontal are designed to give denser coverage area. For example if tree block of size NxN is given, 4N+1 can be used for the prediction. As illustrated in the below diagram there are nine different techniques defined for intra prediction those are
1) Mode 0 : Vertical Comparison
2) Mode 1 : Horizontal Comparison
3) Mode 2 : DC coefficient
4) Mode 3 : Diagonal Down/Left Comparison
5) Mode 4 : Diagonal Down/Right Comparison
6) Mode 5 : Vertical-Right Comparison
7) Mode 6 : Horizontal-Down Comparison
8) Mode 7 : Vertical-Left Comparison
9) Mode 8 : Horizontal-Up Comparison

5.5 Conclusion

From the literature work, it is absorbed that H.265 entropy coding provides higher throughput as compared to H.264 and H.265 does support all the difference types of video application development through varied video coding standards. Still Quad Tree structure of H.265 is very volatile and complexity increases if sparsity technique does not support variability. Commercially it have been proved, that huge bandwidth is saved by the usage of H.265 as compared to H.264 but increasing the frame rate does effect the bitrate efficiency. Many techniques were studied for estimating low complexity encoder approach, almost all of them involved Transform Phase, as this phase deals with conversion of blocks of image pixels into coefficients through which compression of redundant data is achieved. Energy and power is consumed in the transform phase by corresponding coefficients. Coefficient having high amplitudes are considered, rest with zero amplitude are discarded. It does remove redundancy but transform process requires high computation which enhance complexity. Although fast algorithms are introduced giving gain but increasing the computation explained in below transform equation.

$$F(u,v) = \frac{2}{N} G(u)G(v) \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i,j) \cos\left(\frac{(2i+1)u\pi}{2N}\right) \cos\left(\frac{(2j+1)v\pi}{2N}\right) \tag{5.17}$$

$$G(x) = \begin{cases} \frac{1}{\sqrt{2}}, x = 0 \\ 1, otherwise \end{cases} \tag{5.18}$$

Using the above transformation formula for forward and inverse transform requires computation of large number of floating points. For example only for 8x8 pixel block requires 64x64 = 4096 computations to be able to process them for eliminating redundancy. This fact also include that transform is performed at multiple dimensions, first horizontally and vertically separately. Our proposed methodology and encoder architecture eliminates transform phase for reducing complexity. Further we also aim to imply compressive sensing and sparse representation for pixels recognition and random projection cancellation.

Chapter 6

Compressive Sensing based Video Encoding

In this chpater, we first introduce the theory of compressive sensing that forms the basis of our porposed encoding algorithm. Next, we explore the different implementations of compressive sensing techniques and finally we provide an overview of the proposed algorithm and its implementation details.

6.1 Compressive Sensing (CS) for Video Acquisition

The CS theory suggests that a structured (sparse) signal(see Figure 6.1) can be recovered from a small number of linear, incoherent measurements[154, 155]. Consider a signal of interest $x \in R^N$ that is measured as

$$y = \Phi \times x \tag{6.1}$$

where, M is the no of measurements taken per frame, y is the M x 1 compressed matrix, Φ is the M x N measurement matrix and x is the N x 1 pixel array.

6.1.1 Measurement Matrix (Φ) And Transform basis (Ψ)

An important issue of the CS theory is how to construct the measurement matrix. The criteria[156], (our wish list) for the best measurement-matrix has been defined as:

- Near optimal performance .....$O(KlogN)$
- Universality: Φ can be paired with variety of transform matrices (Ψ)
- Low computational complexity

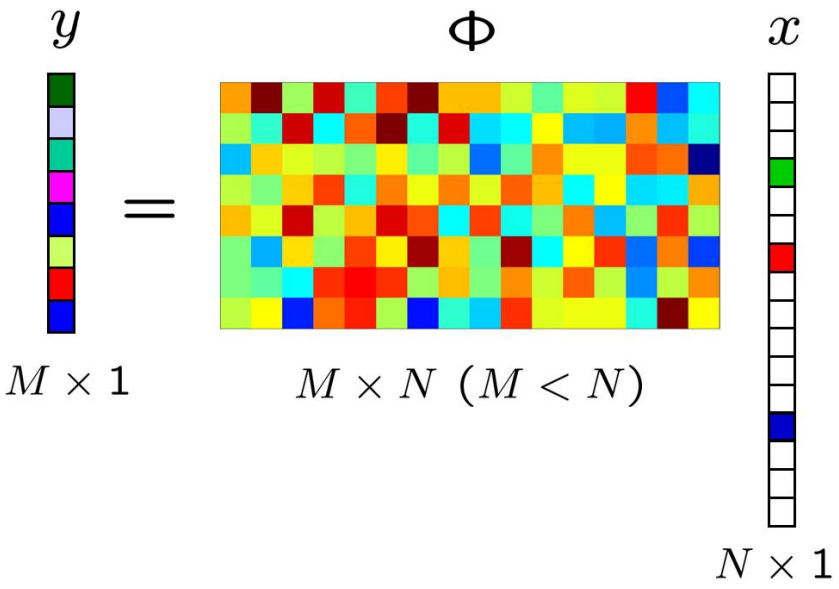


Figure 6.1: The Measurement Matrix.

- Memory efficient
- Hardware friendly i.e. easy to implement.

Romberg, Donoho and many other CS theory experts have stated[157] that a true random measurement matrix gives the best recovery using l1-minimization, provided the original signal x is K-sparse (K<<N) i.e. there are K non-zero values in the vector x. However to insure this (or more precisely enhance sparsity, and hence reduce the number of measurements), we need to find a transform basis (matrix) in which the signal of interest is sparse,

$$y = \Phi x = \Phi \Psi s \tag{6.2}$$

where, Ψ is the transform matrix, s is the K-sparse transform(see Figure 6.2) of x. The Ψ matrix is needed to get the K-sparse x, from the original s matrix which is not K-sparse. In this case[158] we need to take measurements that are incoherent in the Ψ domain i.e. we should be measuring random combinations of basis functions. However in most cases random combinations of pixels behave more or less like random combinations of the basis function(such that they are functionally indistinguishable in practice).

Overall, the transform basis (Ψ) could be divided into three families[157] as follows:

1. The first family are the **random matrices**, including Gaussian random matrix, Bernoulli random matrix, etc. These matrices all obey a certain distribution

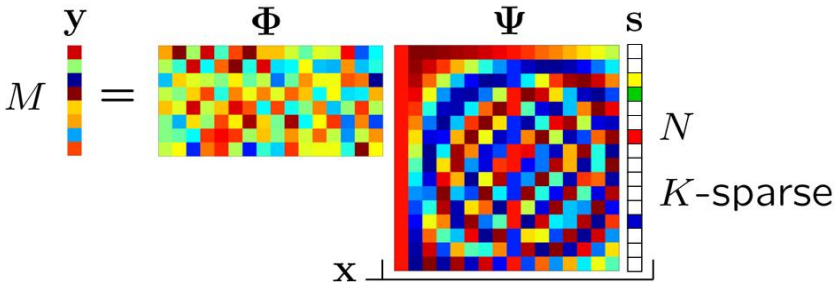


Figure 6.2: The Sparcifying Basis (matrix).

independently.
2. The second family are the matrices generated from the **orthogonal transform matrices**, including Fourier matrix, Hadamard matrix. Such matrices are selecting M rows randomly out of the N N orthogonal matrix, and then normalize each column.
3. The third family are the matrices generated from the **binary orthogonal transform matrices**, including Top Leeds matrix, binary sparse matrix, etc.

6.1.2 The Restricted Isometry Property (RIP)

Any measurement matrix must satisfy the Restricted Isometry Property (RIP), to allow for sparse recovery. Therefore, if the product of the basis matrix and the measurement matrix,
$\Psi \cdot A \cong \Phi\Psi$, satisfies RIP of order k[159], that is

$$(1-\delta_k)\|s\|_{l_2}^2 \leq \|As\|_{l_2}^2 \leq (1+\delta_k)\|s\|_{l_2}^2 \tag{6.3}$$

holds for all k-sparse vectors s for a small isometry constant $0 < \delta_k < 1$, then the sparse coefficient vector s can be accurately recovered via the following linear program:

$$\widehat{s} = \arg\min_{\widetilde{s}} \|\widetilde{s}\|_{l_1} \qquad subject\ to \qquad y = \Phi\Psi s \tag{6.4}$$

Afterward, the signal of interest x can be reconstructed by

$$x = \Psi\widehat{s} \tag{6.5}$$

It has been found[160, 4, 154] that random matrices do satisfy the RIP. The hadamard measurement matrix, when randomized, is a very good option.

6.1.3 The Hadamard Measurement Matrix

The hadamard measurement matrix has been considered by many to be a very good match to the above mentioned criteria, particularly in the context of hardware implementation of CS. Therefore we will construct a Randomized Hadamard

Matrix for our CS measurements[157]. The $N = 2^n$ order Hadamard Matrix $[H_{2^n}]$ is:

$$[H_{2^n}] = [H_2] \times [H_{2^{n-1}}] = \begin{bmatrix} H_{2^{n-1}} & H_{2^{n-1}} \\ H_{2^{n-1}} & -H_{2^{n-1}} \end{bmatrix} \tag{6.6}$$

The Randomized Hadamard measurement matrix is constructed as follows: Firstly generate a N N Hadamard matrix, scramble its columns (optional), and then select M rows out of the matrix randomly to constitute a M N Hadamard measurement matrix.

6.2 Implementation Techniques for Compressive Sensing based Video Acquisition

6.2.1 Optical Implementations

Coded aperture cameras

The random measurement matrix is implemented using a random phase mask[4] placed at the images Fourier plane as depicted in Figure 6.3. The modulated intensity image is then sampled using a low resolution imager to obtain the linear measurement vector. This idea has been demonstrated in IR imaging in which a significant part of the system cost is due to the focal plane array pixel resolution.

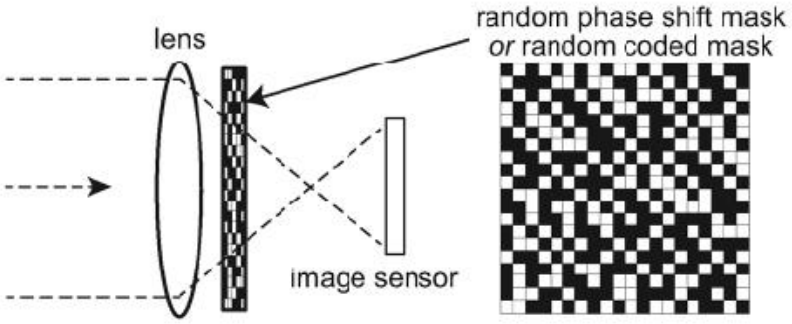


Figure 6.3: Coded aperture cameras[3].

Pros and cons of Coded Aperture cameras are:

1. Implementation of sensing matrix is completely in optical (spatial) domain, so no load on the processor.
2. Measurement is performed in real time.

3. Allows for CS video acquisition.
4. Not block based, so application of block based video encoding (H.264) is not possible.
5. Can be built (similar to the one in Figure 6.4).
6. It suffers from several limitations, especially in visible range imaging:
(i) the optical mask degrades sensitivity,
(ii) precise alignment is needed,
(iii) it is difficult to scale the system resolution.

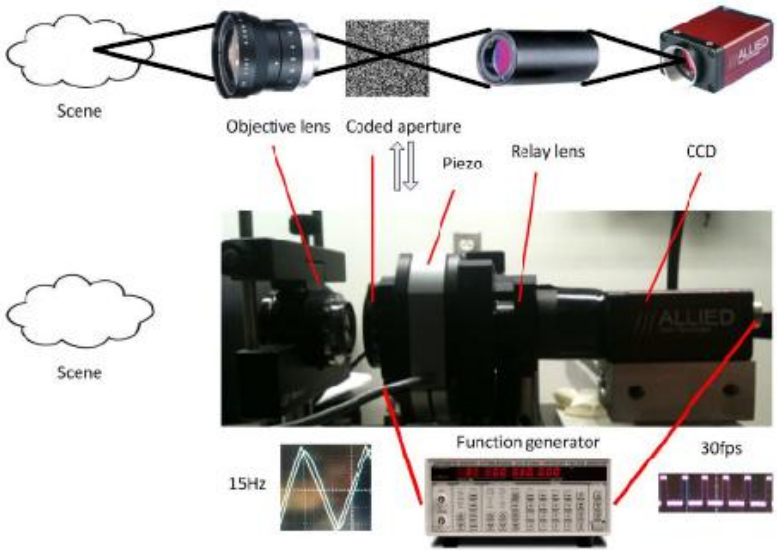


Figure 6.4: Implementation of Coded aperture compressive temporal imaging by, Patrick Llull, Xuejun Liao, Xin Yuan, Jianbo Yang, David Kittle, Lawrence Carin, Guillermo Sapiro, and David J. Brady (www.opticsinfobase.org).

The implementation shown in Figure 6.4 uses a conventional low resolution imager, and the image is captured in a single shot.

Single-pixel camera

The second optical implementation of CS in imaging is the celebrated single pixel camera[4][154] (see Figure 6.5). In this approach, the measurements are acquired sequentially using a single photodiode. The incident light from the scene is reflected off a digital micromirror device (DMD) and the reflected light is collected by a single photodiode. Each mirror can be independently oriented either towards the photodiode (corresponding to a 1) or away from it (corresponding to a 0). To

acquire a measurement, a randomly generated 0/1 vector is used to set the mirror orientation. The output of the pixel represents the sum of the reflected light from the mirrors oriented towards the photodetector.

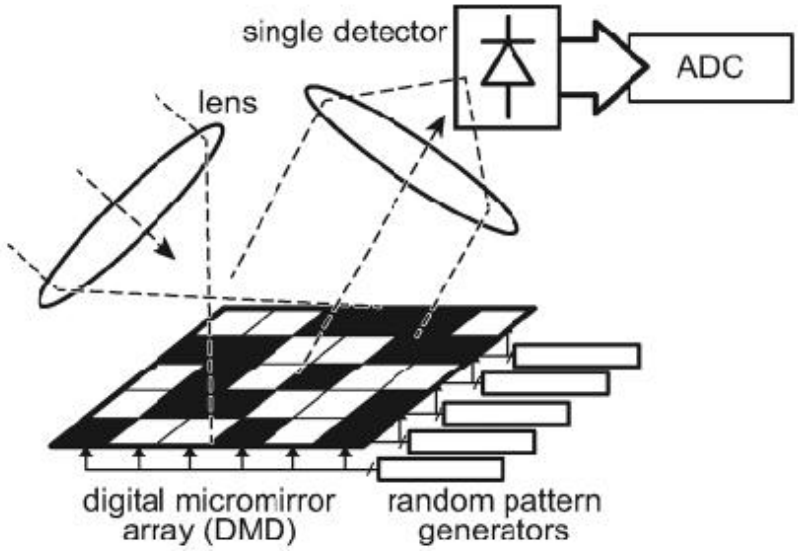


Figure 6.5: Single pixel camera[3].

Pros and cons of Single pixel cameras are:

1. Implementation of sensing matrix will be straight forward since coefficients will be fed directly to the DMD.

2. Measurement is not performed in real time.

3. Does not allow for CS video acquisition.

4. Not block based, so application of block based video encoding (H.264) is not possible.

5. Available in the market.

6. It suffers from several limitations, especially in visible range imaging:

(i) the image is captured using multiple shots, which makes it unsuited for imaging a moving target (the DMDs cannot operate at more than 30kHz, thus limiting the fps to an unacceptable value of 3),

(ii) it is difficult to scale the system resolution,

(iii) reflections from the mirrors result in loss of sensitivity.

The CS work-station(by InView Technology), shown in Figure 6.6, is a very good platform for experimenting with the design of sensing matrices. However it is attractive only when the cost of the photodetector is the dominant component of the total system cost, which is not the case in visible range imaging[4].

Figure 6.6: Compressive sensing work-station by InView Technology.

6.2.2 CMOS Implementations

In-pixel random generators

These sensors have random generator inside each pixel[4, 3]. A flip-flop and control logic to apply random coefficients in two dimensions, simultaneously, to the whole array(see Figure 6.7). Output of each pixel is connected to the corresponding outputs from other pixels, and fed to a trans-impedance amplifier to determine the CS measurement.

Pros and cons of In-pixel random generator based cameras are:
1. Non photo-sensitive elements reduce fill factor, and also the pixel size increases.
2. Measurement is performed in real time.
3. Allows for CS video acquisition.
4. Not block based, so application of block based video encoding (H.264) is not possible.
5. Not available in the market (in IC form).
6. Not implementable at board level because the random coefficients are being applied inside the pixels.

Column-row random selection

Individual access to each pixel is impractical due to layout restrictions. The Column-row random selection architecture[4] is a very good solution. Random coefficients are fed to the rows, then column outputs are randomly combined outside the array to form the CS measurement, which are then passed on to the ADC.

Pros and cons of Single pixel cameras are:
1. Implementation of sensing matrix will become computationally complex because the CS process is divided between the pixel array and the analog processing.
2. Measurement is performed in real time.
3. Allows for CS video acquisition.
4. Not block based, so application of block based video encoding (H.264) is not possible.
5. Not available in the market (in IC form).
6. Implementation at board level is possible thanks to the region of interest (ROI) function available on some modern CMOS sensors.

Block random selection

In this architecture[4] each pixel within a block has a unique path, while the corresponding pixels of each block share the same metal paths.The metal paths are shared among blocks of size 4×4. Random coefficients, V_{s1} through to V_{s16}, can be fed to all the macro-blocks simultaneously (as shown in Figure 6.8 and Figure 6.9). Block switches are used to control the readout so that the blocks can be read out one at a time.

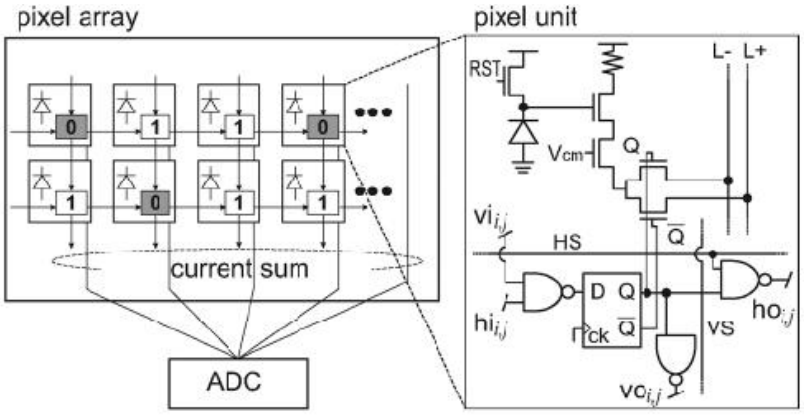


Figure 6.7: In-pixel random generator[3].

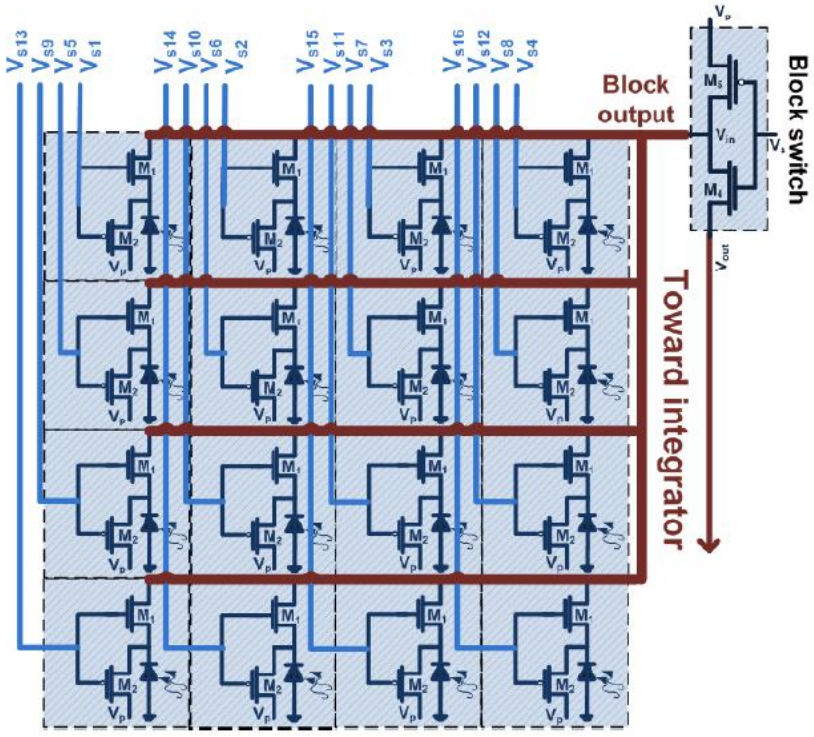


Figure 6.8: A 4x4 block of pixels with their interconnections[4]

The 16 coefficients fed, one coefficient per pixel (as shown in Figure 6.9), are such that the corresponding pixel in each block shares the same coefficient.

Pros and cons of the Block Random Selection method are:
1. Implementation of sensing matrix will be straight forward since coefficients will be fed directly at the sensing stage.
2. With a high pixel rate the measurement can be performed in real time.
3. Allows for CS video acquisition.
4. It is block based so application of H.264 is not possible.
5. Not available in the market (in IC form).
6. Implementation at board level is possible thanks to the region of interest (ROI) function available on some modern CMOS sensors.

6.3 The Proposed CS Video (CSV) Encoder

Using the frame-by-frame, block-by-block CS Video acquisition architecture[159], discussed above, we are able to obtain each frame, Ft, t= 1, 2,, which is divided, at

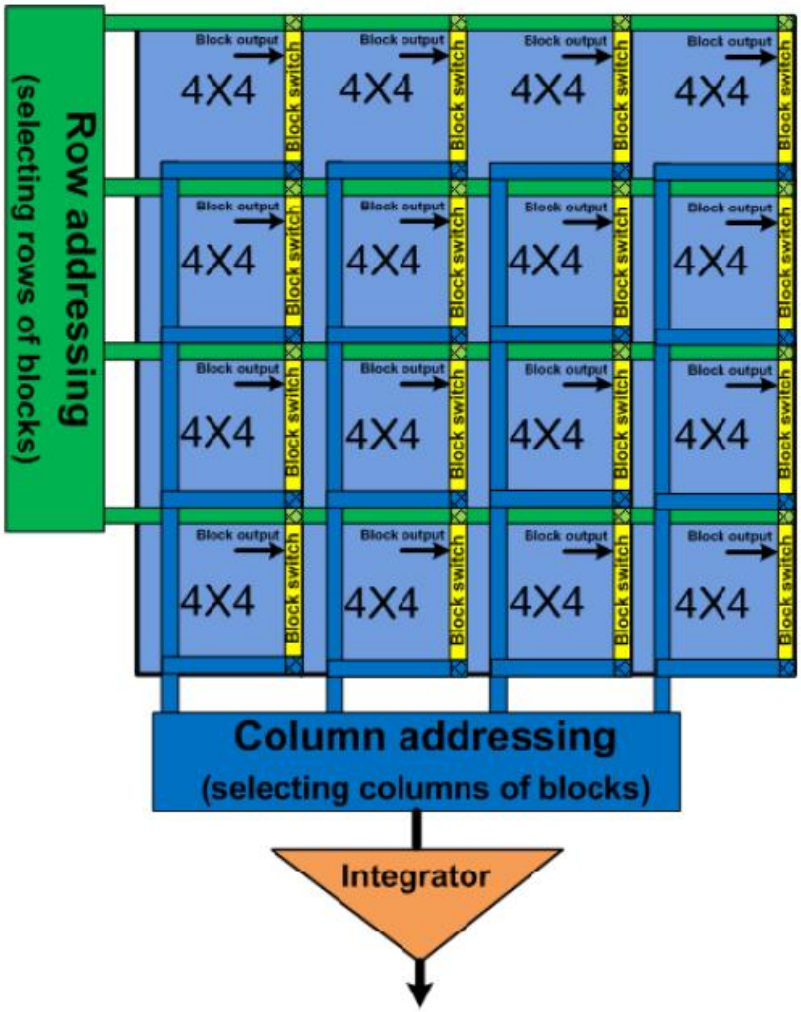


Figure 6.9: 16 blocks with interconnections for block-by-block (BB) read-out[4]

the acquisition stage, into a two dimensional array, I J, of Macro-Blocks (MBs). Each MB consists of pixels, represented as a vectorized column of length N, $x_{t(i,j)} \in R^N$, (i,j) = (1,1),....., (I,J), t = 1,2,.....
Therefore it is a temporal sequence (frames) of spatial raster scan sequences (MBs) as shown below:

$$
\begin{array}{lllllllll}
F_1 = & x_{1(1,1)}, & x_{1(1,2)},\ldots & x_{1(1,J)},\ldots & x_{1(2,1)}, & x_{1(2,2)},\ldots & x_{1(2,J)},\ldots\ldots & x_{1(I,1)}, & x_{1(I,2)}, \\
F_2 = & x_{2(1,1)}, & x_{2(1,2)},\ldots & x_{2(1,J)},\ldots & x_{2(2,1)}, & x_{2(2,2)},\ldots & x_{2(2,J)},\ldots\ldots & x_{2(I,1)}, & x_{2(I,2)}, \\
. & & & & & & & & \\
. & & & & & & & & \\
. & & & & & & & & \\
F_{t-1} = & x_{t-1(1,1)}, & x_{t-1(1,2)},\ldots & x_{t-1(1,J)},\ldots & x_{t-1(2,1)}, & x_{t-1(2,2)},\ldots & x_{t-1(2,J)},\ldots\ldots & x_{t-1(I,1)}, & x_{t-1(I,2} \\
F_t = & x_{t(1,1)}, & x_{t(1,2)},\ldots & x_{t(1,J)},\ldots & x_{t(2,1)}, & x_{t(2,2)},\ldots & x_{t(2,J)},\ldots\ldots & x_{t(I,1)}, & x_{t(I,2)},
\end{array}
\tag{6.7}
$$

As already mentioned CS is performed by projecting $x_{t(i,j)}$ onto a $P \times N$ random measurement matrix Φ,

$$y_{t(i,j)} = x_{t(i,j)}\Phi \tag{6.8}$$

where Φ is a Randomized (scrambled) Hadamard Matrix[156]. The resulting measurement vector $y_{t(i,j)} \in R^P$, is then quantized and passed to the Video Encoder(see Figure 6.10).

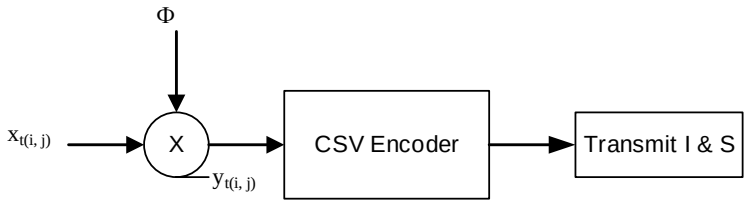


Figure 6.10: The proposed CSV Encoder.

6.3.1 Design considerations

In CS Video encoding, taking the same no of measurements in (each block of) each frame has been proved to be wasteful[155], and better results can be obtained by varying the no of measurements. In our CSV Encoder we will also explore the idea of motion aware, no of measurements on a block by block basis, as will be explained in the following sections. Before moving onto the details of our CSV Encoder, there are a few discussions;

Discussion 1: The optimum raw macro-block (sampling space) size needs to be determined. So the macro-block size will be varied from say, 10 × 10 to 20 × 20, with number of measurements fixed, to see the effect on reconstructed picture quality (on some standard test sequence/picture).

Discussion 2: The optimum no of measurements for each block needs to be determined (not motion aware). So the number of measurements will be varied from say 8 to 32, with macro-block size fixed at say 16×16, to see the effect on reconstructed quality (on some standard test sequence/picture).

Discussion 3: How much gain in terms of bit rate can we get by using a sparcifying basis (such as wavelet, DCT) in addition to the Randomized Hadamard Matrix (which is a very good sensing matrix in terms of its orthogonality, satisfaction of the RIP and ease of application). So we will carry out the above two experiments both with and without the sparcifying basis.

On the basis of these discussions we will be able to determine:

1. Macro-block size
2. Number of measurements for sparse recovery
3. The net. gain on using sparcifying basis (reduced number of measurements/increase in computation)

6.3.2 CSV Encoder design

The Group of Pictures

The Group of Pictures(GoP) is the sequence of frames which repeats over and over again. Our(proposed) CSV Encoder has three frame types, one Start frame(S), m Prediction Frames(P), and n Cluster frames(C). The values of m and n are yet to be determined, i.e. how many frames are needed to make a good prediction and how many frames can be predicted with reasonable accuracy. So for any frame F_t, $t = 1 + m + n$. The overall frame sequence looks like this:

$$S, P_1, P_2, ...P_m C_1, C_2, ..., C_n, S, P_1, P_2, ...P_m C_1, C_2, ..., C_n, S, P_1, P_2, ...P_m C_1, C_2, ..., C_n, ... \tag{6.9}$$

Full frames and Cluster frames Frames are also classified by their acquisition. S frames and P frames are full frames i.e. all the MB's are sensed in a raster scan fashion. C frames are cluster frames i.e. only MB's that fall in clusters are sensed in a broken raster scan fashion.

Figure 6.11 shows the architecture of our CSV Encoder. First there is the Frame Acquisition block, which controls which MB's are to be sensed. It operates either in the full frame mode or the cluster frame mode. In cluster frame mode the MB's to be sensed are given by the Motion Estimation Block(discussed in section; Motion Estimation Block), in the form of a one dimensional array, V_t. V_t contains the exact broken raster scan pattern to be followed at acquisition time.

In the case of full frames, the first frame is intra coded, and the subsequent frames are Inter coded. The Intra Coding block initializes and outputs, a two-dimensional array of Compressed Sensed MB's, S_t, and a two-dimensional array of MB-wise intra prediction pointers, I_t. These S_t and I_t vectors are buffered in the encoder memory for use in Inter Coding(full and cluster) blocks on the arrival of subsequent frames. The Inter Coding(full) block takes as its input, the new full frame(F_t) and the buffered vectors(S_{t-1} and I_{t-1}). The Inter Coding(full) block outputs the updated vectors, S_t and I_t , a two-dimensional vector, R_t, of residual MB's, and the Cluster Pointer vector CP_t. The R_t vector is not buffered and is passed on to the Quantization block, while the CP_t vector is passed to the Motion Estimation block(discussed in section; Motion Estimation Block).
In the case of Cluster frames, the frame(F_t) is passed to the Inter Prediction(cluster) block, which also takes as inputs, the buffered vectors, S_{t-1} and I_{t-1}. This block only outputs the updated vectors, S_t and I_t.

In all cases the vector, I_t, is buffered and forwarded to the Entropy Encode Block and the vector, S_t, is buffered and forwarded to the Δ block, (which also takes as input I_t). The Δ block determines(filters) which MB's in S_t are to be transmitted(ΔS_t). ΔS_t is then passed on to the Quantization block. The Quantization block quantizes all the MB's to be sent and passes the data on to the Entropy Encode block. Finally the entropy Encode block encodes the data and passes it out to the network layer.

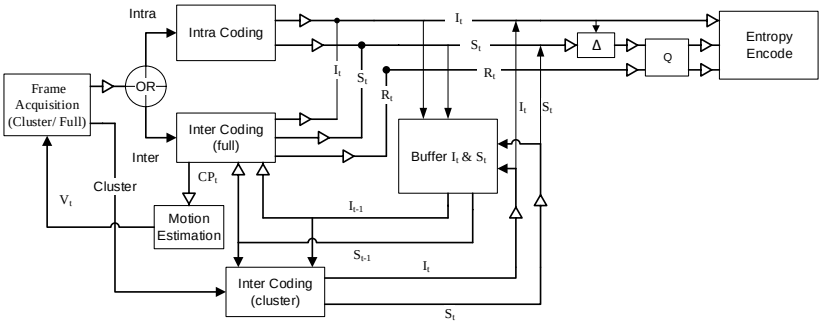


Figure 6.11: CSV Encoder Architecture.

Intra Prediction and the Start Frame(S)

The S frame is a full frame, which comes at the start of each GoP. It is a purely intra prediction frame(very much like the I frame in H.264). The two-dimensional vectors S_t and I_t are initialized(see Figure 6.12). Sum of Absolute Differences (SAD) values of current MB against all previous MBs are calculated, in a simple backward Raster Scan fashion. This is an $O(n^2)$ operation(for the whole frame). If

the minimum SAD value is found to be less than a certain threshold value, b_l , the Intra coding vector I is loaded with a pointer to the corresponding MB, and the current MB is discarded. All the unique MBs of S and the vector I are Entropy Encoded and transmitted.

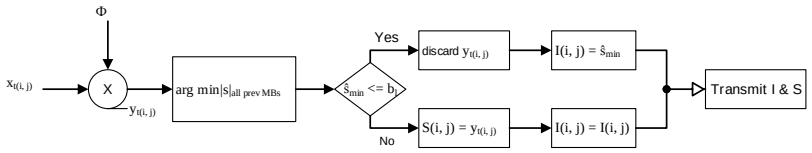


Figure 6.12: Intra prediction in the first frame of each GOP.

Inter Prediction(full) and the Prediction Frames(P)

After the S frame, m P frames are sensed(which are also full frames). They are purely intra prediction frames(very much like the P frame in H.264). The SAD value of each corresponding macro-block of S_t and S_{t-1} will be calculated. Each SAD value is a measure of change, measure of movement within each macro-block. These SAD values will be stored in a two-dimensional vector P (the seme dimensions as S). If a SAD value is less than a lower bound, b_l, then the new macro-block is discarded and the corresponding pointer in I_{t-1} is maintained. If not then the Cluster Update routine is called(see the next section). If a SAD value is less than an upper bound, b_u, then the residual macro-block, $R_{(i,j)}$, is calculated and the original MB in is replaced (corresponding value for the current MB in I vector is also updated). If a SAD value is greater than an upper bound, b_u, then the original MB in S is replaced (corresponding value for the current MB in I vector is also updated). At the end of the frame a mixture of residual and full MBs (together with the updated I vector) will be Entropy Encoded and transmitted(refer to Figure 6.13).

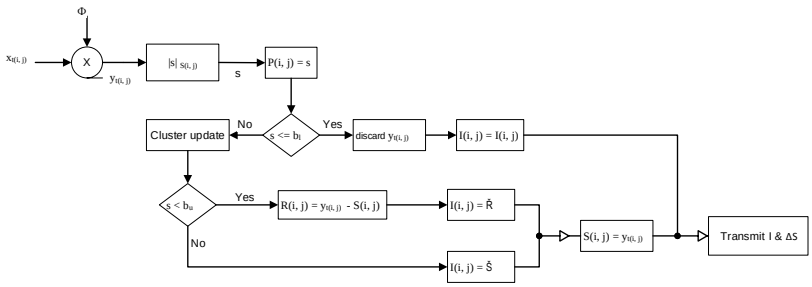


Figure 6.13: Inter prediction between current and previous frames.

Cluster Update

Cluster Update algorithm is called whenever the condition, $s \leq b_l$, is not satisfied(see Figure 6.13). The block diagram in Figure 6.14 is a simplified form of the actual algorithm. On arrival of the first MB (of a P frame) to the Cluster Update block, the cluster variables are initialized(among these is the CP_t). This block gives CP_t as its output to the Motion Estimation block(discussed in next section), at the end of the current P frame. On arrival of each subsequent MB, the algo-

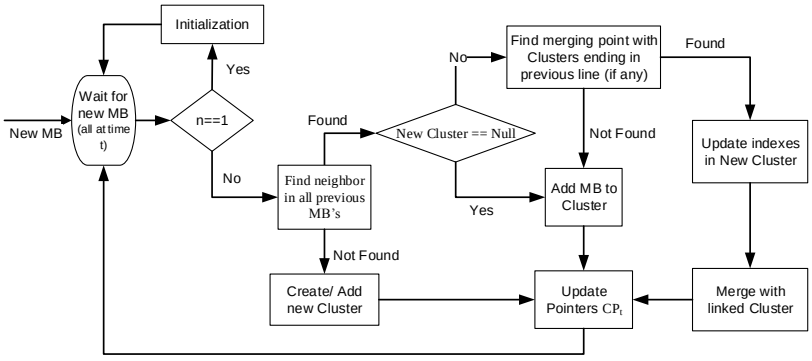


Figure 6.14: Cluster Update Algorithm

rithm searches backward(in raster scan fashion) for a neighbor. If a neighbor is found we check whether we have recently encountered a new cluster. If a neighbor is not found we create a new cluster. Then in the first case we look for the merging point(between new cluster and previous clusters) and make the link. If the merging point is found we merge the new cluster to the linked cluster. If the merging point is not found, the new MB is added to the new cluster. Whether we created a new cluster, or added MB to the new cluster, or merged the new cluster to one of previous clusters, in all cases the CP_t is updated. And as already mentioned the updated CP_t at the end of the frame is passed to the Motion Estimation block.

Motion Estimation

The underlined idea that we are using here is to estimate and predict the movement in different regions(called motion clusters) of the frame(across consecutive frames). The movement can be of many types. In our design we are considering translation and growth(positive and negative) of clusters in the $x-y$ plane. If we determine the mean x and mean y of each cluster and then separately check for change(relative to these mean points) in all four directions, we will get four values(can be positive and negative) for each cluster. These values will be updated on each p frame, until we get a good measure of movement.

The Motion Estimation algorithm is called at the end of each P frame. It takes CP_t as its input from the Inter Prediction(P) block(more precisely from the Cluster Update block). The block diagram in Figure 6.15 is a simplified form of the actual algorithm. On arrival of the first frame(i.e. CP_t), the meanX and meanY values are calculated for each cluster in CP_t . On arrival of each subsequent frame we check to see if the CP_t size is the same as that of CP_{t-1}. If the sizes

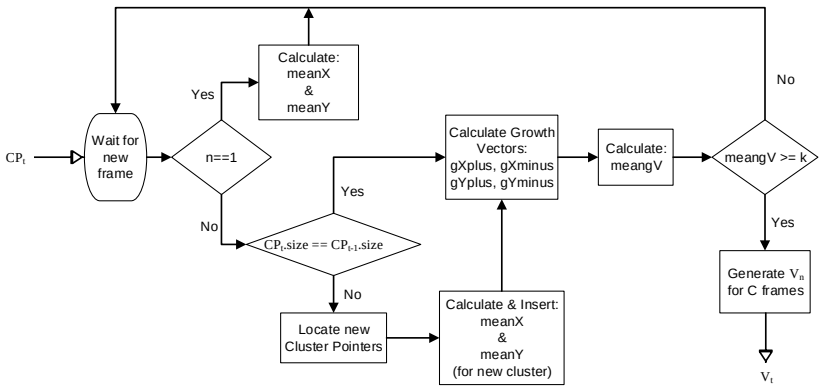


Figure 6.15: The Motion Estimation Algorithm

are not same, then calculate and insert the meanX and meanY values for the new clusters. And in either case we calculate growth vectors(gXplus, gXminus, gYplus, gYminus). Next a mean of growth in the whole frame, meangV, is calculated. The meangV value is an indicator of the overall change in the frame. If the meangV is found to be greater than a certain threshold, k, then we(have enough information) can generate the motion vector, V_t, and start cluster frame acquisition. However if $meangV < k$, then we wait for new P frames until the condition is satisfied. The generated motion vector, V_t, is passed as output to the Frame Acquisition Block.

Inter Prediction(cluster) and the Cluster Frames(C)

Once the Frame Acquisition(cluster/full) block has received the motion vector V_t from the Motion Estimation block, it starts to acquire cluster frames($(C_1, \ C_2, \ C_3, ..., \ C_n)$). The V_t vector is used to extrapolate the MB cluster(s) (see Figure 6.16). Therefore only these predicted cluster frames are compressed sensed. Each MB of C_t is compared (inter predicted) with all the neighboring MBs of S_{t-1} (those that share at least a corner with it), using SAD values. And the minimum of these SAD values, $\hat{s_{min}}$ is retained. If $\hat{s_{min}}$ is less than a threshold, b_l then the new MB($y_{t(i,j)}$) is discarded and the pointer to the MB corresponding to $\hat{s_{min}}$ is saved in corresponding location of I. Otherwise the MB($y_{t(i,j)}$) is saved at $S_{t(i,j)}$ and the I vector remains unchanged. At the end of the frame ΔS and I will be Entropy Encoded and transmitted(refer to Figure 6.15).

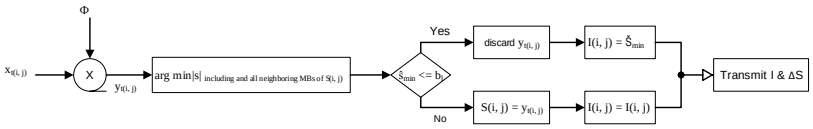


Figure 6.16: Inter prediction between Cluster(s) and the previous frame.

6.3.3 Variable Number of Measurements

Up to this point we are considering that, only a fixed (max) number of measurements are taken for each MB sensed. However the CP_t vectors give us a good idea about the amount of change going on in various areas of the frame (i.e. changes in S_t). So on the basis of these trends an M vector(of same dimensions as the frame) will be updated before the start of each set of frames (at initialization the M will contain the default max value such as 16 on all positions). In this way the number of measurements taken will become adaptable to the amount of change (movement) taking place. Taking lesser measurements will add to bit rate savings.

So as a simple rule, if the SAD values for any MB are below b_l for the last two CP vectors, we could decrement the number of measurements by a factor of 2 (until it comes down to, say, 4). If only one of them is below b_l, we could leave the number of measurements unchanged for that MB. And if both are above b_l, we could increment the number of measurements by a factor of 2 (until it reaches the max value). This is just a raw idea, and a more diverse mechanism may be adopted.

To summarize, by the implementation of CS video acquisition and the CSV Encoder discussed in this report, we can make great gains in terms of reduced encoder complexity, better bit rates, and above all a lot of adaptability.

Chapter 7

Camera Node Architecture and Testbed

This chapter focuses on the specifics of the final prototype, mock deployment scenario and testing matrics. First, the hardware architecture of the camera prototype is discussed which is followed by the description of the testbed comprising five nodes. Finally specific metrices for performance evaluation are discussed.

7.1 The Green Camera Node Architecture

Our Green Camera Node design comprises four major building blocks(see Figure 7.1). At the heart of the node is the micro-controller(μC) with three virtual control units inside it. The Video Acquisition Control Unit(VACU) controls the Video Acquisition Block(VA block). The Energy Management Unit(EMU) controls the Power Source(PS). The Transceiver Control Unit(TCU) controls the 802.11 Transceiver Block(TxRx block). Here we will discuss in some detail the VA block and the PS.

7.1.1 Proposed board level block based implementation of the Video Acquisition Block

Optical implementations suffer from issues like, low resolution, low frame rate, optical alignment and size. Above all block based implementations are not possible. No CS chip is available in the market as yet. Therefore there is no way to feed the sensing matrix coefficients to the sensor array (i.e. at the sensing stage). On the other hand, H.264 encodes the raw (pixel) data in the form of macro blocks. While in CS the pixel data is not available (at the encoder side) so if the whole sensor array is compressively sensed as one unit there will be no way to realistically split the CS raw data (not pixel data) into macro blocks, necessary for H.264 encoding. Therefore it is necessary to perform CS at macro-block level. In

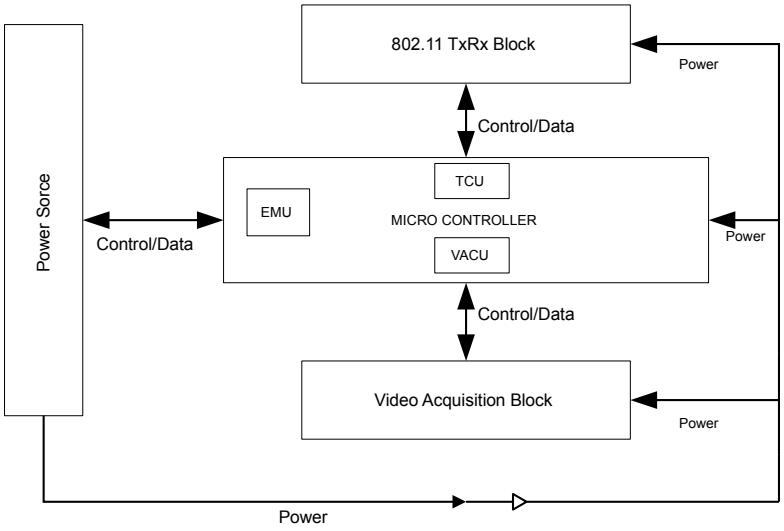


Figure 7.1: Architecture of the Green Camera Node.

this way we will be able to combine CS and a modified version of H.264 (capable of applying H.264 encoding techniques on CS macro blocks rather than ordinary macro blocks)[161, 3, 162].

Architecture of the Video Acquisition Block

In order to achieve the aforementioned requirements we will use the ROI function available on many CMOS cameras[163]. All the macro blocks will be sensed one-by-one as the ROI, and in the form of raw pixel data(Figure 7.2 shows how one MB is processed). In the analog domain, on the arrival of each pixel, we will feed an M long column of the sensing matrix from the microcontroller to allow/suppress the pixel. This column of sensing matrix will be fed to the control lines of the analog switches(such as 18 dg9051)[164], which are ideal for small analog signals of such frequencies. Next the duplicate pixels ($\leq M$) are fed to an integrator which will integrate all the pixels from the current macro block (ROI). Finally these M integrated values are passed on to ADCs giving M measurements(see Figure 7.3). The latency between each macro block (due to SPI communication and related ROI settings) will be the main cause of delay, however we will be able to maintain the real-time requirements of the CS enabled camera node.

The LUPA 300 CMOSS Sensor.

The LUPA 300[163] is a low resolution (640×480) CMOSS sensor with the option of sub sampling, allowing for even lower resolution of 320×240 which is very similar to commercially available WiFi cameras with advanced video codecs (H.264, MJPEG etc.), such as F7D7601 from Belkin. With features like 250fps at

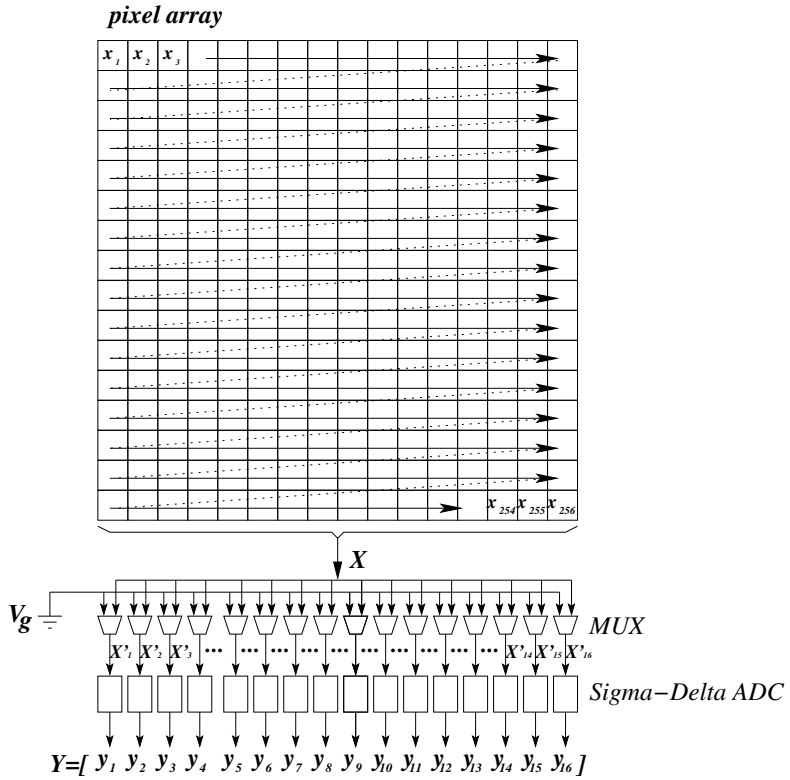


Figure 7.2: Architecture of CS in analog domain[3].

full resolution, Serial Peripheral Interface, programmable readout, windowing and optional analog output, it is an ideal video sensor for the video acquisition on a frame by-frame, block-by-block basis.

The formula for calculating the frame read out time period is as follows:

$$T_{ROI} = FOT + Nr.Lines \times (ROT + Nr.Pixels \times clockperiod) \tag{7.1}$$

This formula will give the time period for only one macro-block. So for a 320×240 frame size and 16×16 macro-block size the formula will become:

$$T_F = 300[FOT + Nr.Lines \times (ROT + Nr.Pixels \times clockperiod)] \tag{7.2}$$

By substituting appropriate values for FOT and ROT (as given in the datasheet) and at nominal clock of 80MHz (0.0125us), the frame time (T_F) is found to be $5.220m \ \ sec.$

To maintain a frame rate of 30fps each frame has about 33m sec (1sec/30 = 33m sec) so we have more than 25m sec to process each frame, and that is

excellent. Its all good provided the read-out time is greater than the integration time (the larger of these will govern the actual frame time period). However this is not always the case, so we will now take a closer look at the LUPA 300 timing.

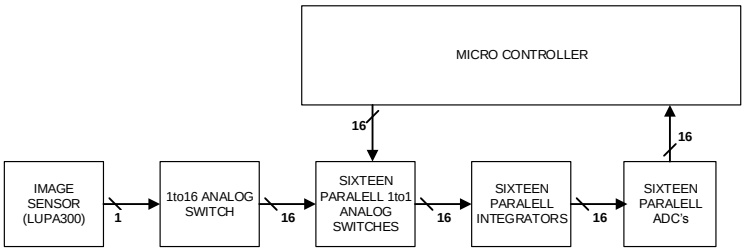


Figure 7.3: The proposed board level implementation using integrators[5] and ADCs.

Integration time and Read-out time

The integration time is the time it takes for the pixel to attain the analog voltage level which represents the light falling on it. This depends mainly on the amount of ambient light present, and that wont always be in our control. In order to enhance the sensor performance, in terms of visual quality, two further integrations, Dual Slope integration (DS) and Triple Slope (TS) integration are also available (the total integration time will be the sum of all three). So there will always be situations when the Integration time will be greater than the Readout time.

The LUPA 300[163] can be operated both in Master and Slave modes. In the Master mode the sensor manages both the Integration and Read-out for the controller, resulting in lesser control and easier implementation. In the Slave mode the Integration time is still managed (primarily) by the sensor, however the Readout time is managed by the controller, resulting in greater control and more timing intensive implementation. At the start of the development, it will be wise to use the sensor in Master mode, and later on it might be of some advantage (in terms of frame time period) to switch to the Slave mode.

To sum up, there are four situations:

1. Master Mode, $T_{int} < T_{read}$
2. Master Mode, $T_{int} > T_{read}$
3. Slave Mode, $T_{int} < T_{read}$
4. Slave Mode, $T_{int} > T_{read}$

There is no point, at this stage to give too many details about the exact timing in each of these four situations. However considering that even when $T_{int} > T_{read}$, it is not so by more than a factor of 2-3, and that we have a margin of 25m sec (as

calculated above), it can be safely said that a Real-Time implementation of the frame by-frame, block-by-block architecture is very practical using the LUPA 300 CMOSS sensor.

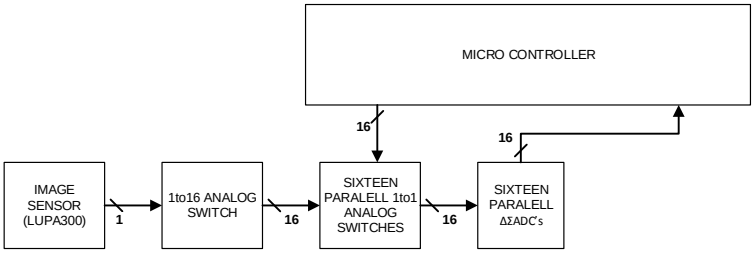


Figure 7.4: The proposed board level implementation using ADCs.

The $\Sigma\Delta$ ADC

In order to simplify our design(see Figure 7.4), we will also explore the use of sigma delta ADCs[165, 166], instead of the integrators and ADCs. Since $\Sigma\Delta$ ADCs are a later technology we will prefer to use them, despite their cost (many times more than ordinary ADCs), because bringing new technology into wider use brings down its price!

The datasheet of AD7760 states:*"The AD7760 is a high performance, 24-bit $\Sigma\Delta$ analog-to-digital converter (ADC). It combines wide input bandwidth and high speed with the benefits of $\Sigma\Delta$ conversion to achieve a performance of 100 dB SNR at 2.5 MSPS, making it ideal for high speed data acquisition. . . . . .*
In addition, the device offers programmable decimation rates, and the digital FIR filter can be adjusted if the default characteristics are not appropriate for the application. The AD7760 is ideal for applications demanding high SNR without a complex front-end signal processing design."

So it is a great option, to simplify the hardware design, and concentrate more on developing new, low complexity CS based encoder design.

Despite the fact that all the details of the design are yet to be figured out, we can quite conclusively say that this analog domain CS implementation is the most realistic so far (available). That is so, at least until the technology gap is closed and the sensing matrix can be fed directly to the sensor array.

7.1.2 Power Source(PS) and the Energy Management Unit

The power source is controlled by the EMU. The system will take power from the battery(see Figure 7.5). The battery is charged from either the DC Source(via a DC plug in available on the node), or the solar Panel(which is an integral part of the node). In this kind of power sources, the battery acts like a source for the "powered components", but acts as a sink to the external source. In a nut-shell, the battery balances the available power from the external source and the demand of the "powered components".

7.2 Testbed

The camera nodes, whose architecture is detailed above would be equipped with customized IEEE 802.11 for video transmission. A layer of functionality will be added on top of the standard 802.11 protocol. The purpose of this layer is to support and control the optimization functions. A number of COTS platforms are available for this purpose. AVS dot11Linux product is once such instance which includes a single board computation platform, IEEE 802.11 source code, development tools, Wi-Fi card and the embedded linux. Parameters of the physical layer will be selected for minimizing energy efficiency. Transmission power, duty cycling, nodes configuration are of the parameters that can optimize the transmission for energy conservation. In the first phase the video sensor will be interfaced with a single board computing device e.g. Raspberry Pi or beagle board that will implement our QUASAR algorithm for video encoding, routing protocol and deliver over customized IEEE 802.11 for transmission. In the second phase, we will develop the integrated green camera node with energy efficient IEEE 802.11 wireless interface after the successful demonstration of our proof-of-the-concept platform.

Performance evaluation of different algorithms is performed both through simulations and hardware implementation. Simulations will be performed in Matlab and Ns-x. Network level performance for large number of nodes will be demonstrated through simulations only. A testbed however, comprising 5 nodes, will be developed for practical network level evaluation. Configuration for this testbed is shown in figure 7.6. For backhauling, any available communication technology will be utilized.

Performance comparisons between H.26x, compressive sensing and QUASAR will be performed through MATLAB simulations. A number of benchmarking videos will tested upon. Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) are standard video quality measurement metrices . These parameters will be measured against all techniques to show the efficiency of our proposed encoding scheme.

IEEE 802.11+ and the green routing protocol will be simulated in NS-2 for the given topology. Performance parameters will include the energy consumption and

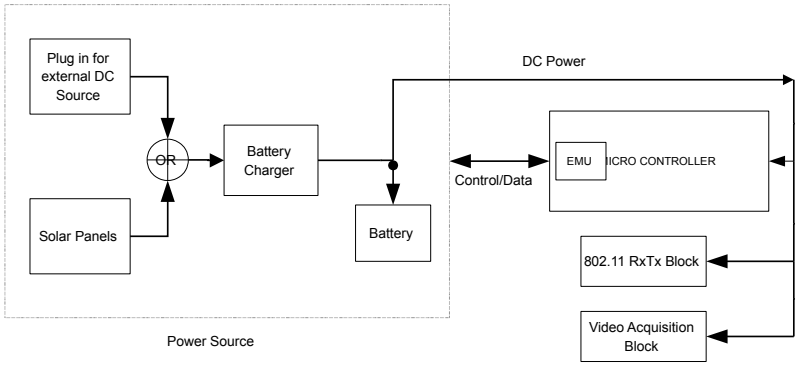


Figure 7.5: The Power Source.

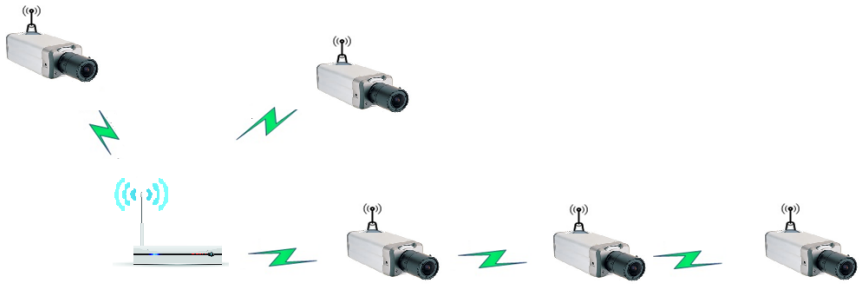


Figure 7.6: Testbed Setup

carbon footprint per bit transmission, throughput, loss rate, delay and jitter. The encoded video data will be generated at three nodes as the leaves of the tree in Figure 10. The scenario will also be demonstrated by using COTS WiFi camera nodes and the performance will be compared with the deployment scenario of green camera test bed deployment. The aforementioned testbed will be developed for a surveillance application. For this application a user friendly GUI shall also be developed. Through this application, the viability of the designed camera nodes and designed protocols for an IoM deployment will be demonstrated.

Bibliography

[1] Daniel Halperin, Ben Greenstein, Anmol Sheth, and David Wetherall. Demystifying 802.11 n power consumption. In *Proceedings of the 2010 international conference on Power aware computing and systems*, page 1. USENIX Association, 2010.

[2] Andres Garcia-Saavedra, Pablo Serrano, Albert Banchs, and Giuseppe Bianchi. Energy consumption anatomy of 802.11 devices and its implication on modeling and design. In *Proceedings of the 8th international conference on Emerging networking experiments and technologies*, pages 169–180. ACM, 2012.

[3] Yusuke Oike and Abbas E.G. "CMOS Image Sensor With Per-Column ADC and Programmable Compressed Sensing". *IEEE JOURNAL OF SOLID-STATE CIRCUITS*, VOL. 48, JANUARY 2013.

[4] Mohammadreza Dadkhah. *"CMOS Image Sensors with Compressive Sensing Acquisition"*. PhD thesis, McMaster University, 2013.

[5] Ray Stata. "Operational Integrators". *Analog Dialog*, APRIL 1967.

[6] Routledge. *The Internet of Things: From RFID to the Next-Generation Pervasive Networked Systems*. Routledge, 2008.

[7] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The internet of things: A survey. *Computer networks*, 54(15):2787–2805, 2010.

[8] P. Friess S. Woelffl H. Sundmaeker, P. Guillemin. *Internet of Things, CERP-IoT - Cluster of European Research Projects on the Internet of Things*. 2010.

[9] J. Belissent. *Getting Clever About Smart Cities: New Opportunities Require New Business Models*. Forrester Research, 2010.

[10] Michele Zorzi, Alexander Gluhak, Sebastian Lange, and Alessandro Bassi. From today's intranet of things to a future internet of things: a wireless-and mobility-related view. *Wireless Communications, IEEE*, 17(6):44–51, 2010.

[11] Olivier Hersent, David Boswarthick, and Omar Elloumi. *The Internet of Things: Key Applications and Protocols*. John Wiley & Sons, 2011.

[12] J. Mogul H. Frystyk L. Masinter P. Leach R. Fielding, J. Gettys and T. Berners-Lee. *The Internet of Things: Key Applications and Protocols*. HyperText Transfer Protocol - HTTP/1.1, RFC 2616, Internet Engineering Task Force RFC 2616, June 1999. Available online: http://www.rfc-editor.org/rfc/rfc2616.txt.

[13] J. Postel. *Internet Protocol, RFC 791, Internet Engineering Task Force RFC 791*. September 1981.

[14] *Transmission Control Protocol, RFC 793, Internet Engineering Task Force RFC 793*. September 1981. Available online: http://www.rfc-editor.org/rfc/rfc793.txt.

[15] Yinghui Huang and Guanyu Li. Descriptive models for internet of things. In *Intelligent Control and Information Processing (ICICIP), 2010 International Conference on*, pages 483–486. IEEE, 2010.

[16] George Lawton. Machine-to-machine technology gears up for growth. *Computer*, 37(9):12–15, 2004.

[17] ETSI TS 102 689 v1.1.1. *Machine-to-Machine communications (M2M): M2M service requirements*. Aug. 2010.

[18] Jean-Philippe Vasseur and Adam Dunkels. *Interconnecting smart objects with ip The next internet*. Morgan Kaufmann, 2010.

[19] Adam Dunkels and Jean-Philippe Vasseur. Ip for smart objects, september 2008. *IPSO Alliance White Paper*, 1.

[20] *IEEE 802.11n-2009 Standard for Information technology, Local and metropolitan area networks, Specific requirements, Part 11: Wireless LAN Medium Access Control (MAC)and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput*. IEEE.

[21] Byoung Hoon Jung, Hu Jin, and Dan Keun Sung. Adaptive transmission power control and rate selection scheme for maximizing energy efficiency of ieee 802.11 stations. In *Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on*, pages 266–271. IEEE, 2012.

[22] *Z-Wave. Available online: http://www.z-wave.com/.*

[23] K West, G West, and K Hall. Wtrs wireless sensor network technology trends report. Technical report, Technical report, Western Technology Research Solutions, 2008.

[24] *Wireless Connectivity Market Data.* ABI Research, 2011.

[25] *3D-CineCast,"http://3dcinecast.blogspot.com/2011 10 01 archive.html".*

[26] Gary J Sullivan, Jens Ohm, Woo-Jin Han, and Thomas Wiegand. Overview of the high efficiency video coding (hevc) standard. *Circuits and Systems for Video Technology, IEEE Transactions on*, 22(12):1649–1668, 2012.

[27] *VP9 Google Video Coding, "http://www.webmproject.org/vp9/".* Jan 2014.

[28] *Daala Video Coding, "http://people.xiph.org/ xiph-mont/demo/daala/demo4.shtml".* Feb 2014.

[29] E. Fleisch. *What is the Internet of Things? - An Economic Perspective.* Auto-ID Labs, Tech. Rep, 2010.

[30] *European Research Cluster on Internet of Things (IERC).* Internet of Things - Pan European Research and Innovation Vision, IERC, Available online: http://www.internet-of-things-research.eu/documents.html, 2011.

[31] Luca Mainetti, Luigi Patrono, and Antonio Vilei. Evolution of wireless sensor networks towards the internet of things: A survey. In *Software, Telecommunications and Computer Networks (SoftCOM), 2011 19th International Conference on*, pages 1–6. IEEE, 2011.

[32] *IEEE standard 802.15.4, Part. 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks(LR-WPANs), Standard for Information Technology.* 2006.

[33] *ZigBee Alliance, Available online: www.zigbee.org.*

[34] *D. Networks, Available online: ww.dustnetworks.com.*

[35] K Pister and Lance Doherty. Tsmp: Time synchronized mesh protocol. *IASTED Distributed Sensor Networks*, pages 391–398, 2008.

[36] *HART Communication Protocol and Fundation , Available online: http://www.hartcomm2.org.*

[37] G. Montenegro N. Kushalnagar and C. Schumacher. *IPv6 over Low- Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals, Internet Engineering Task Force RFC 4919.* August 2007.

[38] A. Brandt J. Hui R. Kelsey P. Levis K. Pister R. Struik J. P. Vasseur T. Winter, P. Thubert and R. Alexander. *RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks, RFC 6550, Internet Engineering Task Force RFC 6550.* March 2012.

[39] C. Bormann Z. Shelby, K. Hartke and B. Frank. *Constrained Application Protocol (CoAP), IETF CoRE Working Group.* February 2011.

[40] IEEE. *802.15.4e-2012: IEEE Standard for Local and Metropolitan Area Networks - Part 15.4: Low-Rate Wireless Personal Area Networks (LRWPANs) Amendment 1: MAC Sublayer.* April 2012.

[41] Andrew Tinka, Thomas Watteyne, and Kris Pister. A decentralized scheduling algorithm for time synchronized channel hopping. In *Ad Hoc Networks*, pages 201–216. Springer, 2010.

[42] S. Berson S. Herzog R. Braden, L. Zhang and S. Jamin. *Resource ReSerVation Protocol (RSVP) - Version 1 Functional Speci
cation, RFC 2205, Internet Engineering Task Force RFC 2205.* September 1997.

[43] IETF. *MPLS-TP Internet Drafts and RFCs.*

[44] Thomas Watteyne, Ankur Mehta, and Kris Pister. Reliability through frequency diversity: why channel hopping makes sense. In *Proceedings of the 6th ACM symposium on Performance evaluation of wireless ad hoc, sensor, and ubiquitous networks*, pages 116–123. ACM, 2009.

[45] Branko Kerkez, Thomas Watteyne, Mario Magliocco, Steven Glaser, and Kris Pister. Feasibility analysis of controller design for adaptive channel hopping. In *Proceedings of the Fourth International ICST Conference on Performance Evaluation Methodologies and Tools*, page 76. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2009.

[46] Jonathan W Hui and David E Culler. Extending ip to low-power, wireless personal area networks. *Internet Computing, IEEE*, 12(4):37–45, 2008.

[47] JeongGil Ko, Andreas Terzis, Stephen Dawson-Haggerty, David E Culler, Jonathan W Hui, and Philip Levis. Connecting low-power and lossy networks to the internet. *Communications Magazine, IEEE*, 49(4):96–101, 2011.

[48] J. Hui G. Montenegro, N. Kushalnagar and D. Culler. *Transmission of IPv6 Packets over IEEE 802.15.4 Networks, RFC 4944, Internet Engineering Task Force RFC 4944.* September 2007.

[49] J. Hui and P. Thubert. *Compression Format for IPv6 Datagrams over IEEE 802.15.4- Based Networks, RFC 6282, Internet Engineering Task Force RFC 6282.* September 2011.

[50] M. Crawford. *Transmission of IPv6 Packets over Ethernet Networks, RFC 2464, Internet Engineering Task Force RFC 2464.* December 1998.

[51] G. Fairhurst D. Grossman R. Ludwig J.Mahdavi G. Montenegro J. Touch P. Karn, C. Bormann and L. Wood. *Advice for Internet Subnetwork Designers, RFC 3819, Internet Engineering Task Force RFC 3819.* July 2004.

[52] Mikael Degermark. *IP header compression, RFC 2507.* 1999.

[53] Zach Shelby and Carsten Bormann. *6LoWPAN: The wireless embedded Internet*, volume 43. John Wiley & Sons, 2011.

[54] Jonathan W Hui and David E Culler. Ipv6 in low-power wireless networks. *Proceedings of the IEEE*, 98(11):1865–1878, 2010.

[55] Paul Bertrand, SAS Watteco, and Cedric Chauvenet. Rpl: The ip routing protocol designed for low power and lossy networks. 2011.

[56] J. Martocci. *Building Automation Routing Requirements in Low-Power and Lossy Net- works, RFC 5867, Internet Engineering Task Force RFC 5867.* June 2010.

[57] T. Winter M. Dohler, T. Watteyne and D. Barthel. *Routing Requirements for Urban Low-Power and Lossy Networks, RFC 5548, Internet Engineering Task Force RFC 5548.* May 2009.

[58] K. Pister N. Dejean J. P. Vasseur, M. Kim and D. Barthe. *Routing Metrics Used for Path Calculation in Low-Power and Lossy Networks, RFC 6552, Internet Engineering Task Force RFC 6552.* March 2012.

[59] G. Fairhurst and L. Wood. *Advice to Link Designers on Link Automatic Repeat reQuest (ARQ), RFC 3366, Internet Engineering Task Force RFC 3366.* August 2002.

[60] J. Postel. *User Datagram Protocol, RFC 768, Internet Engineering Task Force RFC 768.* August 1980.

[61] *Constrained RESTful Environments (core). IETF Working Group, Available online: http://www.ietf.org/dyn/wg/charter/core-charter.html.*

[62] Carsten Bormann, Angelo P Castellani, and Zach Shelby. Coap: An application protocol for billions of tiny internet nodes. *IEEE Internet Computing*, 16(2), 2012.

[63] George Adam, Vaggelis Kapoulas, Christos Bouras, Georgios Kioumourtzis, Apostolos Gkamas, and Nikos Tavoularis. Performance evaluation of routing protocols for multimedia transmission over mobile ad hoc networks. In *Wireless and Mobile Networking Conference (WMNC), 2011 4th Joint IFIP*, pages 1–6. IEEE, 2011.

[64] S. Pink L-E. Jonsson Ed L-A. Larzon, M. Degermark. *The Lightweight User Datagram Protocol (UDP-Lite), Available: http://tools.ietf.org/html/rfc3828.*

[65] S. Floyd E. Kohler, M. Handley. *Datagram Congestion Control Protocol (DCCP), Available: http://www.ietf.org/rfc/rfc4340.txt.*

[66] Dominic Lenton. Speaking of wi-fi. *IEE Review*, 49(7):44–47, 2003.

[67] IEEE. *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications. IEEE Std 802.11-1999, 1999.*

[68] Eldad Perahia and Robert Stacey. Next generation wireless lans. *Throughput, robustness and reliability in 802.11 n*, 2008.

[69] H Alan and H Chi-Yu. Medium access control. In *Computer Communications and Networks*, pages 35–50, 2010.

[70] Osama Aboul-Magd. *Wireless Local Area Networks Quality of Service: An Engineering Perspective.* IEEE, 2009.

[71] J-S Lee. Performance evaluation of ieee 802.15. 4 for low-rate wireless personal area networks. *Consumer Electronics, IEEE Transactions on*, 52(3):742–749, 2006.

[72] Pablo Serrano, Albert Banchs, and Arturo Azcorra. A throughput and delay model for ieee 802.11 e edca under non saturation. *Wireless Personal Communications*, 43(2):467–479, 2007.

[73] Hua Zhu, Ming Li, Imrich Chlamtac, and B Prabhakaran. A survey of quality of service in ieee 802.11 networks. *Wireless Communications, IEEE*, 11(4):6–14, 2004.

[74] Xavier Pérez-Costa and Daniel Camps-Mur. Ieee 802.11 e qos and power saving features overview and analysis of combined performance [accepted from open call]. *Wireless Communications, IEEE*, 17(4):88–96, 2010.

[75] Konstantinos Pelechrinis, Theodoros Salonidis, Henrik Lundgren, and Nitin Vaidya. Experimental characterization of 802.11 n link quality at high rates. In *Proceedings of the fifth ACM international workshop on Wireless network testbeds, experimental evaluation and characterization*, pages 39–46. ACM, 2010.

[76] Robert C Daniels, Constantine M Caramanis, and Robert W Heath. Adaptation in convolutionally coded mimo-ofdm wireless systems through supervised learning and snr ordering. *Vehicular Technology, IEEE Transactions on*, 59(1):114–126, 2010.

[77] Pankaj Bhagawat, Rajballav Dash, and Gwan Choi. Array like runtime reconfigurable mimo detectors for 802.11 n wlan: a design case study. In *Proceedings of the 2009 Asia and South Pacific Design Automation Conference*, pages 751–756. IEEE Press, 2009.

[78] Richard van Nee. 802.11 n: The global wireless lan standard. In *Globalization of Mobile and Wireless Communications*, pages 103–118. Springer, 2011.

[79] Haifeng Zheng, Guotai Chen, and Lun Yu. Video transmission over ieee 802.11 n wlan with adaptive aggregation scheme. In *Broadband Multimedia Systems and Broadcasting (BMSB), 2010 IEEE International Symposium on*, pages 1–5. IEEE, 2010.

[80] Emna Charfi, Lamia Chaari, and Lotfi Kamoun. Phy/mac enhancements and qos mechanisms for very high throughput wlans: a survey. *IEEE Communications Surveys & Tutorials*, 2013.

[81] *Wireless Gigabit Alliance. De
ning the future of multi-gigabit wireless communications.* White Paper, pp. 1-5, 2010.

[82] Robert Stacey, Eldad Perahia, A Stephens, et al. Specification framework for tgac. *doc.: IEEE802*, pages 11–09, 2010.

[83] Eldad Perahia and Michelle X Gong. Gigabit wireless lans: an overview of ieee 802.11 ac and 802.11 ad. *ACM SIGMOBILE Mobile Computing and Communications Review*, 15(3):23–33, 2011.

[84] Chunhui Zhu, Youngsoo Kim, Osama Aboul-Magd, and Chiu Ngo. Multi-user support in next generation wireless lan. In *Consumer Communications and Networking Conference (CCNC), 2011 IEEE*, pages 1120–1121. IEEE, 2011.

[85] Valeria Baiamonte and C-F Chiasserini. Saving energy during channel contention in 802.11 wlans. *Mobile Networks and Applications*, 11(2):287–296, 2006.

[86] Sylwia Romaszko and Chris Blondia. Neighbour-aware, collision avoidance mac protocol (ncmac) for mobile ad hoc networks. In *Wireless Communication Systems, 2006. ISWCS'06. 3rd International Symposium on*, pages 322–326. IEEE, 2006.

[87] Xiaodong Wang, Jun Yin, and Dharma P Agrawal. Effects of contention window and packet size on the energy efficiency of wireless local area network. In *Wireless Communications and Networking Conference, 2005 IEEE*, volume 1, pages 94–99. IEEE, 2005.

[88] Xiaodong Wang, Jun Yin, and Dharma P Agrawal. Analysis and optimization of the energy efficiency in the 802.11 dcf. *Mobile networks and applications*, 11(2):279–286, 2006.

[89] Mustafa Ergen and Pravin Varaiya. Decomposition of energy consumption in ieee 802.11. In *Communications, 2007. ICC'07. IEEE International Conference on*, pages 403–408. IEEE, 2007.

[90] Subir Biswas and Samir Datta. Reducing overhearing energy in 802.11 networks by low-power interface idling. In *Performance, Computing, and Communications, 2004 IEEE International Conference on*, pages 695–700. IEEE, 2004.

[91] Mathieu Lacage, Mohammad Hossein Manshaei, and Thierry Turletti. Ieee 802.11 rate adaptation: a practical approach. In *Proceedings of the 7th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems*, pages 126–134. ACM, 2004.

[92] Ivaylo Haratcherev, Koen Langendoen, Reginald Lagendijk, and Henk Sips. Hybrid rate control for ieee 802.11. In *Proceedings of the second international workshop on Mobility management & wireless access protocols*, pages 10–18. ACM, 2004.

[93] Ad Kamerman and Leo Monteban. Wavelan®-ii: a high-performance wireless lan for the unlicensed band. *Bell Labs technical journal*, 2(3):118–133, 1997.

[94] Starsky HY Wong, Hao Yang, Songwu Lu, and Vaduvur Bharghavan. Robust rate adaptation for 802.11 wireless networks. In *Proceedings of the 12th annual international conference on Mobile computing and networking*, pages 146–157. ACM, 2006.

[95] Shuguang Cui, Andrea J Goldsmith, and Ahmad Bahai. Energy-constrained modulation optimization. *Wireless Communications, IEEE Transactions on*, 4(5):2349–2360, 2005.

[96] Byoung Hoon Jung, Hu Jin, and Dan Keun Sung. Adaptive transmission power control and rate selection scheme for maximizing energy efficiency of ieee 802.11 stations. In *Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on*, pages 266–271. IEEE, 2012.

[97] Paul Lettieri and Mani B Srivastava. Adaptive frame length control for improving wireless link throughput, range, and energy efficiency. In *INFOCOM'98. Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 564–571. IEEE, 1998.

[98] Naomi Ramos, Debashis Panigrahi, and Sujit Dey. Energy-efficient link adaptations in ieee 802.11 b wireless lan. In *Proc. of International Conference on Wireless and Optical Communications*. Citeseer, 2003.

[99] Dinesh Rajan and Christian Poellabauer. Adaptive fragmentation for latency control and energy management in wireless real-time environments. In *Wireless Algorithms, Systems and Applications, 2007. WASA 2007. International Conference on*, pages 158–168. IEEE, 2007.

[100] Marcelo M Carvalho, Cintia B Margi, Katia Obraczka, and JJ Garcia-Luna-Aceves. Modeling energy consumption in single-hop ieee 802.11 ad hoc networks. In *Computer Communications and Networks, 2004. ICCCN 2004. Proceedings. 13th International Conference on*, pages 367–372. IEEE, 2004.

[101] Daji Qiao, Sunghyun Choi, Amit Jain, and Kang G Shin. Miser: an optimal low-energy transmission strategy for ieee 802.11 a/h. In *Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 161–175. ACM, 2003.

[102] Jyh-Cheng Chen and Kai-Wen Cheng. Edca/ca: Enhancement of ieee 802.11 e edca by contention adaption for energy efficiency. *Wireless Communications, IEEE Transactions on*, 7(8):2866–2870, 2008.

[103] Raffaele Bruno, Marco Conti, and Enrico Gregori. Optimization of efficiency and energy consumption in p-persistent csma-based wireless lans. *Mobile Computing, IEEE Transactions on*, 1(1):10–31, 2002.

[104] Francesco Ivan Di Piazza, Stefano Mangione, and Ilenia Tinnirello. On the effects of transmit power control on the energy consumption of wifi network cards. In *Quality of Service in Heterogeneous Networks*, pages 463–475. Springer, 2009.

[105] Andres Garcia-Saavedra, Pablo Serrano, Albert Banchs, and Matthias Hollick. Energy-efficient fair channel access for ieee 802.11 wlans. In *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2011 IEEE International Symposium on a*, pages 1–9. IEEE, 2011.

[106] Yu Xiao, Yong Cui, Petri Savolainen, Matti Siekkinen, An Wang, Liu Yang, Antti Ylä-Jääski, and Sasu Tarkoma. Modeling energy consumption of data transmission over wi-fi. 2013.

[107] Andrew Rice and Simon Hay. Measuring mobile phone energy consumption for 802.11 wireless networking. *Pervasive and Mobile Computing*, 6(6):593–606, 2010.

[108] Mustafa Ergen and Pravin Varaiya. Decomposition of energy consumption in ieee 802.11. In *Communications, 2007. ICC'07. IEEE International Conference on*, pages 403–408. IEEE, 2007.

[109] Siavash Alamouti. A simple transmit diversity technique for wireless communications. *Selected Areas in Communications, IEEE Journal on*, 16(8):1451–1458, 1998.

[110] Vahid Tarokh, Hamid Jafarkhani, and A. Robert Calderbank. Space-time block codes from orthogonal designs. *Information Theory, IEEE Transactions on*, 45(5):1456–1467, 1999.

[111] Intel wireless wifi link drivers for linux. http://intellinuxwireless.org/.

[112] IEEE 802 LAN/MAN Standards Committee et al. Ceet, the power of wireless cloud: An analysis of the energy consumption of wireless cloud, centre for energy-efficient telecom- munications, tech. rep., april 2013, online: http://www.ceet.unimelb. edu.au/pdfs/ceet white paper wireless cloud.pdf. *IEEE, http://standards. ieee. org/getieee802/download/802.11 n-2009. pdf*, 2009.

[113] J-M Tarascon. Key challenges in future li-battery research. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1923):3227–3241, 2010.

[114] Geoffrey Ye Li, Zhikun Xu, Cong Xiong, Chenyang Yang, Shunqing Zhang, Yan Chen, and Shugong Xu. Energy-efficient wireless communications: tutorial, survey, and open issues. *Wireless Communications, IEEE*, 18(6):28–35, 2011.

[115] Shiao-Li Tsao and Chung-Huei Huang. A survey of energy efficient mac protocols for ieee 802.11 wlan. *Computer Communications*, 34(1):54–67, 2011.

[116] Pi-Cheng Hsiu, Chun-Han Lin, and Cheng-Kang Hsieh. Dynamic backlight scaling optimization for mobile streaming applications. In *Proceedings of the 17th IEEE/ACM international symposium on low-power electronics and design*, pages 309–314. IEEE Press, 2011.

[117] Fang-wei Li, Ya-qing Zhang, and Liang-wei Li. Enhanced discontinuous reception mechanism for power saving in td-lte. In *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on*, volume 9, pages 682–686. IEEE, 2010.

[118] Justin Manweiler and Romit Roy Choudhury. Avoiding the rush hours: Wifi energy management via traffic isolation. In *Proceedings of the 9th international conference on Mobile systems, applications, and services*, pages 253–266. ACM, 2011.

[119] Kostas Pentikousis. In search of energy-efficient mobile networking. *Communications Magazine, IEEE*, 48(1):95–103, 2010.

[120] Selvadurai Selvakennedy and Sukunesan Sinnappan. An energy-efficient clustering algorithm for multihop data gathering in wireless sensor networks. *Journal of Computers*, 1(1):40–47, 2006.

[121] Surendar Chandra and Amin Vahdat. Application-specific network management for energy-aware streaming of popular multimedia formats. In *USENIX Annual Technical Conference, General Track*, pages 329–342, 2002.

[122] Pavlos Petoumenos, Georgia Psychou, Stefanos Kaxiras, Juan Manuel Cebrian Gonzalez, and Juan Luis Aragon. Mlp-aware instruction queue resizing: The key to power-efficient performance. In *Architecture of Computing Systems-ARCS 2010*, pages 113–125. Springer, 2010.

[123] *Low Energy Consumption for Networks of Future*. Available Online, http://www.econet-project.eu/, 2011.

[124] Henrik Petander. Energy-aware network selection using traffic estimation. In *Proceedings of the 1st ACM workshop on Mobile internet through cellular networks*, pages 55–60. ACM, 2009.

[125] Daniel Camps-Mur, Manil Dev Gomony, Xavier Pérez-Costa, and Sebastià Sallent-Ribes. Leveraging 802.11 n frame aggregation to enhance qos and power consumption in wi-fi networks. *Computer Networks*, 56(12):2896–2911, 2012.

[126] Janet Adams and Gabriel-Miro Muntean. Adaptive-buffer power save mechanism for mobile multimedia streaming. In *Communications, 2007. ICC'07. IEEE International Conference on*, pages 4548–4553. IEEE, 2007.

[127] Rajesh Palit, Kshirasagar Naik, and Ajit Singh. Impact of packet aggregation on energy consumption in smartphones. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2011 7th International*, pages 589–594. IEEE, 2011.

[128] ME Khanouche, FS Ouada, and S Ouguigui. Energy-aware clustering and sensor scheduling coverage maintenance for wireless sensor networks. In *Green Computing and Communications (GreenCom), 2012 IEEE International Conference on*, pages 171–178. IEEE, 2012.

[129] Mohammadjavad Abbasi, Muhammad Shafie Bin Abd Latiff, and Hassan Chizari. An overview of distributed energy-efficient topology control for wireless ad hoc networks. *Mathematical Problems in Engineering*, 2013, 2013.

[130] Eunhwa Kim, Jeoungpil Ryu, and Kijun Han. An energy efficient network topology configuration scheme for sensor networks. In *Advanced Web and Network Technologies, and Applications*, pages 297–305. Springer, 2006.

[131] J. Buron A. Brandt and G. Porcu. *Home Automation Routing Requirements in Low- Power and Lossy Networks, RFC 5826, Internet Engineering Task Force RFC 5826.* April 2011.

[132] K. S. J. Pister and P. Thubert. *Industrial Routing Requirements in Low-Power and Lossy Networks, RFC 5673, Internet Engineering Task Force RFC 5673.* October 2009.

[133] Shawmin Lei and Shijun Sun. Video compression for raw rgb format using residual color transform, March 18 2005. US Patent App. 10/907,082.

[134] Chun-Su Park. Spatial and interlayer hybrid intra-prediction for h. 264/svc video. *Optical Engineering*, 52(7):071503–071503, 2013.

[135] Iain E Richardson. *The H. 264 advanced video compression standard.* John Wiley & Sons, 2011.

[136] Xiang Gen Xia, Charles Boncelet, and Gonzalo Arce. Wavelet transform based watermark for digital images. *Optics Express*, 3(12):497–511, 1998.

[137] *Transform Coding,"http://www.deetc.isel.ipl.pt/analisedesinai/ccd/docs/JKieffer notes13.pdf".* Jan 2014.

[138] Jinjia Zhou, Dajiang Zhou, and Satoshi Goto. Alternating asymmetric search range assignment for bidirectional motion estimation in h. 265/hevc and h. 264/avc. *Journal of Visual Communication and Image Representation*, 2014.

[139] Dan Grois, Detlev Marpe, Amit Mulayoff, Benaya Itzhaky, and Ofer Hadar. Performance comparison of h. 265/mpeg-hevc, vp9, and h. 264/mpeg-avc encoders.

[140] Malavika Bhaskaranand and Jerry D Gibson. Global motion compensation and spectral entropy bit allocation for low complexity video coding. In *Communications (ICC), 2012 IEEE International Conference on*, pages 2043–2047. IEEE, 2012.

[141] Liang Zhao, Li Zhang, Siwei Ma, and Debin Zhao. Fast mode decision algorithm for intra prediction in hevc. In *Visual Communications and Image Processing (VCIP), 2011 IEEE*, pages 1–4. IEEE, 2011.

[142] Roopa.M. Chitra.M1. An optimized distributed video coding using turbo codes and zero motion skip encoder strategy, issn:2278-2834 volume 1, issue2: (may-june 2012), pp 10-14.

[143] Anne Aaron and Bernd Girod. Wyner-ziv video coding with low-encoder complexity. In *Proc. Picture Coding Symposium, PCS 2004*, 2004.

[144] Heming Sun, Dajiang Zhou, and Satoshi Goto. A low-complexity hevc intra prediction algorithm based on level and mode filtering. In *Multimedia and Expo (ICME), 2012 IEEE International Conference on*, pages 1085–1090. IEEE, 2012.

[145] Xianguo Zhang, YongHong Tian, Tiejun Huang, and Wen Gao. Low-complexity and high-efficiency background modeling for surveillance video coding. In *VCIP*, pages 1–6. IEEE, 2012.

[146] Linux news, "http://linuxfr.org/news/h-265-est-finalise", 2013.

[147] Gary J Sullivan and Stephen J Estrop. Methods and systems for start code emulation prevention and data stuffing, 2009. US Patent 7,505,485.

[148] Ntu university, "http://www.cmlab.csie.ntu.edu.tw/cml/dsp/training/coding/jpeg/jpeg/encoder.h 2013.

[149] *Motion-Prediction,"http://wiki.multimedia.cx/index.php?title=Motion Prediction"*. 2007.

[150] Hevc video coding standard, "http://xevc.net/" jan 2014.

[151] Frank Bossen, Benjamin Bross, Karsten Suhring, and David Flynn. Hevc complexity and implementation analysis. *Circuits and Systems for Video Technology, IEEE Transactions on*, 22(12):1685–1696, 2012.

[152] Vivienne Sze and Madhukar Budagavi. A comparison of cabac throughput for hevc/h. 265 vs. avc/h. 264. In *Signal Processing Systems (SiPS), 2013 IEEE Workshop on*, pages 165–170. IEEE, 2013.

[153] Code:sequoia,"http://codesequoia.wordpress.com/2009/10/18/h-264-stream-structure/", 2009.

[154] Marco F. Duarte, Mark A. Davenport, Dharmpal Takhar, Jason N. Laska, Ting Sun, Kevin F. Kelly, and Richard G. Baraniuk. "Single-Pixel Imaging via Compressive Sampling". Technical report, Digital Vision, 2008.

[155] M. Salman Asif, Felix Fernandes, and Justin Romberg. "Low-Complexity Video Compression and Compressive Sensing". In *Asilomar*, 2013.

[156] Lu Gan, Thong T. Do, and Trac D. Tran. "Fast Ccomressive Imaging using Scrambled Block Hadamard Ensemble". In *Proc. EUSIPCO*, 2008.

[157] Rui Sun, Honglin Zhao, and Hao Xu. "The Application of Improved Hadamard Measurement Matrix in Compressed Sensing". In *International Conference on Systems and Informatics (ICSAI 2012)*, 2012.

[158] Justin Romberg. "Imaging via Compressive Sampling". Technical report, Digital Vision, MARCH 2008.

[159] Ying Liu, Ming Li, and Dimitris A. Pados. "Motion-Aware Decoding of Compressed-Sensed Video". *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY*, 23(3), MARCH 2013.

[160] Scott Pudlewski, Tommaso Melodia, and Arvind Prasanna. "Compressed-Sensing-Enabled Video Streaming for Wireless Multimedia Sensor Networks". In *IEEE TRANSACTIONS ON MOBILE COMPUTING*, volume 11, JUNE 2012.

[161] Saad Qaisar, Rana Muhammad Bilal, Wafa Iqbal, Muqaddas Naureen, and Sungyoung Lee. "Compressive Sensing: From Theory to Applications, A Survey". Technical report, National University of Sciences and Technology Islam- abad, Pakistan., 2001.

[162] J.W. Chen, C.Y. Kao, and Y.L. Lin. "Introduction to H.264 Advanced Video Coding". In *Conference on Asia South Pacific design automation, New York, NY, USA*, 2006.

[163] ON Semiconductor. *"LUPA300 CMOS Image Sensor"*, 2013.

[164] "Precision 8-Ch/Dual 4-Ch/Triple 2-Ch Low Voltage Analog Switches/Multiplexers", MAY 2011.

[165] "Sigm-Delta ADC's and DAC's".

[166] Analog Devices. *"Sigma-Delta ADC with On-Chip Buffer"*, 2006.