MEASURING VITAL SIGNS USING SMART PHONES

Vikram Chandrasekaran, B.E.

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

December 2010

APPROVED:

Ram Dantu, Major Professor
Philip H. Sweany, Committee Member
Parthasarathy Guturu, Committee Member
Ian Parberry, Interim Chair of the
Department of Computer Science
and Engineering
Costas Tsatsoulis, Dean of the College of
Engineering
James D. Meernik, Acting Dean of the
Robert B. Toulouse School of
Graduate Studies

Chandrasekaran, Vikram. Measuring Vital Signs Using Smart Phones. Master of Science (Computer Science), December 2010, 130 pp., 23 tables, 77 illustrations, references, 56 titles.

Smart phones today have become increasingly popular with the general public for its diverse abilities like navigation, social networking, and multimedia facilities to name a few. These phones are equipped with high end processors, high resolution cameras, built-in sensors like accelerometer, orientation-sensor, light-sensor, and much more. According to comScore survey, 25.3% of US adults use smart phones in their daily lives. Motivated by the capability of smart phones and their extensive usage, I focused on utilizing them for bio-medical applications.

In this thesis, I present a new application for a smart phone to quantify the vital signs such as heart rate, respiratory rate and blood pressure with the help of its built-in sensors. Using the camera and a microphone, I have shown how the blood pressure and heart rate can be determined for a subject. People sometimes encounter minor situations like fainting or fatal accidents like car crash at unexpected times and places. It would be useful to have a device which can measure all vital signs in such an event.

The second part of this thesis demonstrates a new mode of communication for next generation 9-1-1 calls. In this new architecture, the call-taker will be able to control the multimedia elements in the phone from a remote location. This would help the call-taker or first responder to have a better control over the situation. Transmission of the vital signs measured using the smart phone can be a life saver in critical situations. In today’s voice oriented 9-1-1 calls, the dispatcher first collects critical information (e.g., location, call-back number) from caller, and assesses the situation. Meanwhile, the

dispatchers constantly face a "60-second dilemma"; *i.e.,* within 60 seconds, they need to make a complicated but important decision, whether to dispatch and, if so, what to dispatch. The dispatchers often feel that they lack sufficient information to make a confident dispatch decision. This remote-media-control described in this system will be able to facilitate information acquisition and decision-making in emergency situations within the 60-second response window in 9-1-1 calls using new multimedia technologies.

ACKNOWLEDGMENTS

This work is partially supported by the National Science Foundation under grants CNS-0751205 and CNS-0821736.

I would like to acknowledge Dr. Ram Dantu to have supported me through my Masters. I would also like to thank the Network Security Lab members and medical professionals: Dr. Chandrasekaran Ganapathy, Dr. Yuvraj Radhakrishnan and Dr. Amudhan for assisting and aiding me throughout my works.

CONTENTS

ACKNOWLEDGEMENTS iii

LIST OF TABLES ix

LIST OF FIGURES xi

CHAPTER 1. BIOMEDICAL APPLICATIONS OF SMART PHONES 1

1.1. Introduction 1

1.2. The Vital Signs 2

CHAPTER 2. ESTIMATION OF HEART RATE WITH MOBILE CAMERA 6

2.1. Introduction 6

2.2. Medical Protocol 6

2.3. Brief Description of Proposed Method 7

2.4. Detailed Methodology 7

2.5. Implementation of Methodology 9

2.6. Heart Rate Accuracy with Proposed Method 11

2.7. Window Time Calibration for Accuracy Improvement 12

2.8. Impact on Heart Rate with Age and Physical Intensity 14

2.9. Conclusion 16

CHAPTER 3. ESTIMATION OF RESPIRATORY RATE WITH MOBILE SENSORS 18

3.1. Introduction 18

3.2. Medical Protocol 18

3.3. Brief Description of Proposed Method 18
3.4. Detailed Methodology 19
3.5. Respiratory Pattern from Oxygen Saturation 23
3.5.1. Discussion on Respiratory Patterns 24
3.6. Conclusion 25

CHAPTER 4. CUFFLESS BLOOD PRESSURE ESTIMATION WITH MOBILE PHONE 27
4.1. Introduction 27
4.2. Terminologies Used 28
4.3. Traditional Protocols 29
4.4. White-coat Syndrome 30
4.5. Brief Description of Proposed Methods 30
4.6. Cuff Based Estimation 31
4.6.1. Detailed Methodology 31
4.7. Cuffless Estimation 36
4.7.1. Detailed Methodology 37
4.7.2. Accuracy with Cuffless Estimation 40
4.7.3. Sensitivity Analysis for Cuffless Estimation Method 42
4.7.4. Degree of Significance in Identification of Two Pressure Groups 47
4.8. Cuff-less Differential Estimation 49
4.8.1. Detailed Methodology 49
4.8.1.1. Localization of Heart Beat and Pulse 50
4.8.2. Inter Working of Components for Blood Pressure Estimation 52
4.8.3. Bluetooth based Clock Synchronization 55
4.8.3.1. Empirical Framework 56

4.8.3.2. Master-Slave Message Exchange Mechanism 57
4.8.3.3. Slave Device synchronization to Master Clock 58
4.8.4. Accuracy with Differential Estimation 61
4.8.5. Impact of Age, Height and Weight on Systolic Pressure 65
4.8.6. Mobile hardware Performance Analysis 67
4.8.6.1. Bluetooth performance 67
4.8.6.2. Microphone Performance 70
4.8.7. Limitations of Methodology 71
4.9. Conclusion 72

CHAPTER 5. REMOTE MEDIA CONTROL 74
5.1. Introduction 74
5.2. Need for Media Control 75
5.2.1. Description of Communication Components 75
5.2.2. Description of Mobile Client 79
5.2.3. Message Exchange between Mobile and PC Client 80
5.2.4. Control Message Transfer Process 82
5.3. Image Transfer 83
5.3.1. Description 84
5.3.2. Camera Specification 85
5.3.3. FTP Based Image Transfer using SIP 85
5.3.4. Event Flow for SIP Assisted FTP Transfer 87
5.3.5. Impact of Media Features on Image Quality 89
5.3.5.1. Perceived Quality to Image Resolution 89
5.3.5.2. Perceived Quality to Brightness, Contrast and Sharpness 90
5.3.5.3. Perceived Quality to Camera Exposure 93

5.4. System Performance 94
5.4.1. Architecture of Utilization Monitor 94
5.4.2. Network Utilization of SIPUA 95
5.4.3. CPU Utilization of SIPUA 99
5.4.4. Call Quality Analysis of System 100
5.4.4.1. Voice Quality under Different Microphone Modes 103
5.5. Conclusion 104

CHAPTER 6. SOCIO-TECHNICAL ASPECTS OF VIDEO PHONES 105
6.1. Introduction 105
6.2. Social Impacts 105
6.3. Social Issues 109
6.3.1. Display of Mood and Emotions 110
6.3.2. Trustworthiness 110
6.3.3. Unexpected Video Clips 110
6.3.4. No Control over Surroundings 111
6.4. Technical Issues 111
6.4.1. PC-Based Video Conferencing 111
6.4.2. Comparison of Social Issues in Three Electronic Communication Systems 112
6.4.3. Size of the Video Screen 112
6.4.4. Perception of Motion and Distance 115
6.4.4.1. Distance from the Camera 115
6.4.4.2. Brightness of the Environment 117
6.4.5. Change in Bandwidth Due to Human Activities 119
6.5. Conclusion 120

CHAPTER 7. CONCLUSION 122

7.1. Challenges and Limitations 123
7.2. Future Work 124
REFERENCES 126

LIST OF TABLES

1.1 Heart Rate and Respiratory Rate for Different Ages 3

1.2 Categories of Blood Pressure 4

2.1 Accuracy of Heart Rate 11

2.2 Projected Error in Methodology. 12

2.3 Calibrated Results of Heart Rate 15

4.1 Accuracy of Pulse Pressure 34

4.2 Accuracy of Cuff-Based Blood Pressure Estimation 36

4.3 Test Subject Information 41

4.4 Accuracy of Cuffless Blood Pressure Estimation 41

4.5 Estimated mean arterial pressure for two individuals 48

4.6 Heart Valve Location over Chest 52

4.7 Accuracy of Blood Pressure for Different Peak Types 62

4.8 Accuracy of Systolic Pressure in Differential Estimation 63

4.9 Subject information 67

4.10 Bluetooth Synchronization Performance 69

5.1 User and Dispatcher Clients Feature Comparison 79

5.2 Media Control Ranges. 88

5.3 File Size Vs Lighting. 93

5.4 Microphone Quality on Different Modes 102

5.5 Call Quality on Different Modes 103

6.1 PC Video Conferencing vs Video Phone 112

6.2 Social Issues of Electronic Communications 113

6.3 Activity Identification with Bitrate 120

LIST OF FIGURES

2.1 Locations for Pulse Detection 7
2.2 Method of Measurement 8
2.3 Architecture 8
2.4 Analysis Region 9
2.5 Filtering Data 10
2.6 Data Analysis 10
2.7 Calibration of Heart Rate 13
2.8 Heart Rate Sensitivity for Age and Work Intensity 16

3.1 Respiratory Rate Setup 19
3.2 Architecture of Respiratory Rate Detection 20
3.3 Acceleration Graph under Different Scenarios 20
3.4 Acceleration Graph from Different Regions on Chest 21
3.5 Acceleration Graph for Different Modes of breathing 22
3.6 Acceleration Detection on All Axis 22
3.7 Respiratory Disorder Diagnosis 23
3.8 Respiratory Pattern from Oxygen Saturation 25
3.9 Respiratory Pattern from Video 26

4.1 Blood Pressure from Sphygmomanometry 30
4.2 Test Setup for Cuff-Based Estimation 32

4.3 Relationship between Cuff Inflation and Blood Pressure 32
4.4 Estimation of Impedance Factor 35
4.5 Estimating Ejection Time 38
4.6 Correlation of Ejection Time 39
4.7 Estimation of Reference Resistance 40
4.8 Sensitivity Analysis on Ejection Time 43
4.9 Sensitivity Analysis on Cardiac Output 44
4.10 Sensitivity Analysis on Pulse Pressure 45
4.11 Sensitivity Analysis on Resistance 46
4.12 T-Test for Two Pressure Groups 47
4.13 Test Setup for Differential Blood Pressure Estimation 50
4.14 Localizing Heart Sounds on Chest 51
4.15 Differential Blood Pressure Estimation Working 53
4.16 Peak Identification 55
4.17 Vascular Transit Time and Systolic Pressure relation 56
4.18 Bluetooth Synchronization Test Setup 57
4.19 Synchronization Architecture 58
4.20 Synchronization Event Flow 59
4.21 Client Clock Synchronization 60
4.22 Types of Peaks 61
4.23 Continuous Pressure Estimation 64
4.24 Accuracy of Systolic Pressure with Number of Data Points 65
4.25 Sensitivity Analysis of Differential Estimation 66
4.26 Structure of Data Packet 69

4.27 Quality of Microphone on Different Phones 71

5.1 System Design of Remote Media Control 76

5.2 Prototype of Dispatcher Client 77

5.3 Prototype of User Client 78

5.4 Message Exchange between Clients 80

5.5 Architecture of Mobile Client 81

5.6 Event Flow between Clients 84

5.7 Architecture of SIP Assisted Image transfer 86

5.8 Image Transmission Module 87

5.9 Message Flow in Image Transfer 88

5.10 File Size Vs Image resolution 90

5.11 Transmission Time Vs Image Resolution 90

5.12 File Size Vs Zoom factor 91

5.13 File Size Vs Brightness 92

5.14 File Size Vs Contrast 92

5.15 File Size Vs Sharpness 93

5.16 Architecture of Utilization Monitor 95

5.17 Utilization Monitor Test Setup 96

5.18 Network Utilization on Idle 97

5.19 Network Utilization on Voice Call 97

5.20 Network Utilization on Moving Scene Video Call 98

5.21 Network Utilization on Static Scene Video Call 98

5.22 CPU Utilization on Idle 99

5.23 CPU Utilization on Video Call 100

5.24 Call Quality Analysis 101

6.1 Diagnosis with Video Phones 108

6.2 Video Conferencing by All Age groups 108

6.3 Inappropriate Presentation on Video Phone 110

6.4 Bitrate Vs Quality 114

6.5 Distance Vs Motion Detection 116

6.6 Distance Vs Quality 116

6.7 Motion Detection Vs Brightness 117

6.8 Quality Vs Brightness 118

6.9 Blocking Vs Brightness 119

CHAPTER 1

BIOMEDICAL APPLICATIONS OF SMART PHONES

1.1. Introduction

With the advent of internet, a lot has changed in people's lives. Users can stay in the house and do almost any kind of activity like shopping, movies, entertainment, physical exercise, without physically being at the appropriate place. In short, we can tell that our lives are made simpler, faster and efficient, if we leave behind the negative aspects of it. In the field of telecommunication, video phones have a wide scope in bringing people together with face-to-face communication. Further more, video transmission from mobile phone as a video call enhances the mobility of users. In case of a 911 situation, this will allow people to make conversations a lot more understandable even under chaos. The responders could arrive at a decision more quickly, when viewing the live video feed of the person. The caller's information about the physical condition and vital signs during that time will enhance the responder's ability to assist them in the best possible way.

This thesis concentrates mainly on two vivid regions: one is the diagnosis of vital signs of the body and how a mobile phone can help in biotelemetry, and the other is the means of sending all the data through remote connection in case of a life threatening emergency situation. I have done extensive studies in vital signs identification from a person and propose new methods of using mobile phone's sensors for quantifying the vital sign. These methods utilize the accelerometer, video camera in tandem with a LED flash and microphone sensor to detect and deliver a value for the needs. In this chapter, I provide a general discussion about the vital signs and the need for its diagnosis.

1.2. The Vital Signs

Vital signs are the most basic functions that can be measured from a person;they indicate their physical condition and wellness. When the measurements tend to move away from normal, an abnormality in the physical status can be inferred. Most medical conditions can be diagnosed through vital signs and confirmed with the help of special tests. Each vital sign is measured differently with the use of specialized equipments. These equipments are not handy and do not come in miniature sizes for portability. Hence I introduce the concept of converting a mobile phone, which people use in their day to day life into a vital sign diagnosing tool. There are four vital signs which are standard in most medical settings:

- Pulse rate
- Respiratory rate
- Blood pressure
- Temperature

Pulse rate (HR) is the rate at which the heart beats, measured either in the wrist or neck given by beats per minute. The pulse rate is influenced by the expansion of the arterial wall for every beat. The most prominent spots for the pulses are wrist (Radial artery), neck(Carotid artery), inside of the elbow (Brachial artery), behind the knee(Popliteal artery) and ankle joint (Posterior tibial artery) [1]. The pulse rate varies with age and also depends on the physical and psychological effects on the body. Higher pulse rate indicates the presence of abnormality in the body and can also be caused by other reasons such as anxiety, anger, excitement, emotion, heart disorders, asthma, a large meal and so on. The pulse rate of an individual can help in determining various problems within the body, but it cannot be used solely to diagnose an abnormality. Pulse rate is just a basic tool for diagnosis and hence can be used only for primary diagnosis.

Respiratory rate (RR) is the number of breaths a person takes within a certain amount of time or more formally, defined as the number of chest movements involving inspiration and expiration per unit time. The RR is measured in units of breaths per minute. It is measured by counting the number of breaths(number of times the chest rise) for a minute, usually when the person is at rest. Respiratory rates will increase as the demand for oxygen increases; it also increases due to illness, intensive physical activity, etc. The average RR reported for a healthy adult at rest is usually given as 12 breaths per minute (12/60 Hz) [2] and the estimates vary between 12-20 breaths per minute, whereas the respiratory rate is higher in the case of young adults, children and babies. As people age, breathing rate declines. In slow rates, more accurate readings are obtained by counting the number of breaths over a full minute. Table 1.1 [3] shows the heart rate and respiratory rate at varying ages showing a gradual decline in the rate with age.

Table 1.1. *Heart Rate and Respiratory Rate for Different Ages*

Age	Heart Rate (beats/min)	Respiratory Rate(breaths/min)
Newborn	100-160	30-50
0-5 months	90-150	25-40
6-12 months	80-140	20-30
1-3 years	80-130	20-30
3-5 years	80-120	20-30
6-10 years	70-110	15-30
11-14 years	60-105	12-20
14+ years	60-100	12-20

Blood pressure (BP) is a force exerted by blood on the walls of arteries, veins and the chambers of the heart. Blood pressure is one the most important vital signs and the body maintains it by interacting with the volume of blood and the force of contraction of the heart. During each heartbeat, BP varies between a maximum pressure called systolic pressure and a minimum pressure called diastolic pressure. It is measured on the inside of an elbow at the brachial artery, which is the upper arm's major blood vessel that carries blood away from the heart. A person's BP is usually expressed in terms of the systolic pressure and diastolic pressure values. An average healthy adult's pressure values read 120 mmHg during the systole and 80 mmHg during diastole. Pumping Rate, blood volume, resistance, viscosity, etc. are some of the factors which affect the blood pressure of a person. Due to various reasons, the average blood pressure differs from each individual. The pressure values are categorized into five major divisions. Table 1.2 shows the categories of people in accordance to their blood pressure range.

Table 1.2. *Categories of Blood Pressure*

Category	Systolic (mmHg)	Diastolic (mmHg)
Hypotension	$<$90	$<$60
Normal	90-120	60-80
Prehypertension	121-139	or 81-89
Stage 1 Hypertension	140-159	or 90-99
Stage 2 Hypertension	$\geq$ 160	or = 100

Temperature (T) is one of the other important vital signs. There is no direct way of measuring the person's temperature from the mobile device as of now. Mobile phone manufacturers have started incorporating onboard eco temperature sensors to mobile phones. It is just a matter of time before the temperature of surroundings and a human body can be measured using the phone.

This thesis focuses on using only a mobile phone as a device for measuring heart rate, respiratory rate and blood pressure of a person. The remaining chapters of this thesis are as follows. I have discussed the methodologies and results obtained from different experiments for measuring heart rate, respiratory rate and blood pressure in detail under Chapters 2, 3 and 4 respectively. Chapter 5 throws light into remote media control which is the main application for communication. An integral part of this chapter is the discussion of a new system for controlling media elements such as microphone, speaker, camera, etc., remotely. Chapter 6 discusses some of the preliminary works that have been published relating to social and technical aspects of video phones. Finally the thesis is concluded in Chapter 7.

CHAPTER 2

ESTIMATION OF HEART RATE WITH MOBILE CAMERA

2.1. Introduction

A pulse is a sudden burst of blood to the circulatory system when the walls of the heart contract. Heart rate or pulse rate is defined as the number of heart beats or pulses in a minute. The human heart comprises the atrium and the ventricles, which coordinate to form a complete pumping action. Approximately 2000 gallons of blood is pumped by the heart every day.A Heart beat cycle consists of two components, namely systole and diastole. Systole occurs when there is an electrical impulse generated by the Sinoatrial(SA) Node, triggering the heart to contract. Diastole occurs when the heart is relaxed. Systole and diastole alternate each other to produce a heart beat. The heart rate is not just about how fast the heart is beating; it is a regulatory mechanism for delivering oxygen to the muscles to keep up the demand.

2.2. Medical Protocol

Acoustically, the heart rate is measured by listening to the heart beats, which are amplified through the use of a stethoscope. Usually the number of beats for a small interval of time, say 10 seconds, is observed and obtained for a minute by multiplying with 6. In the same way, the pulse felt at the wrist and neck can be measured and directly related to the heart rate. Figure 2.1 shows the regions where the pulse can be felt clearly for measurement. A more precise method of determining pulse rate involves the use of an electrocardiography(ECG or EKG), pulse oximetry, etc. Shelley [4] discussed about the effectiveness of pulse oximetry in the detection of pulse even under noisy conditions where the use of stethoscope is hopeless. There are many commercial heart rate monitors available in the market which use two tiny

Figure 2.1. *Prominent place for pulse detection.*

electrode strips to find the heart rate, the same way an ECG works. These electrodes are generally attached to some fitness gear or costume, displaying the measurements on a screen. ECG uses the electrical activity of the heart over a period of time, measured through the electrodes connected to the skin. These electrodes induce a tiny current of a few μA into the body and detect electrical changes caused by the heart during each heart beat. These changes are captured, amplified and delivered as an output.

2.3. Brief Description of Proposed Method

Even with the presence of many technologies for finding the heart rate of a person, only a few of them are accurate to a certain degree. In this section I proposed a new model for heart rate estimation, which works on the concept of Photo plethysmography (PPG), without using the wavelength of light for analysis. Most mobile phones in today's market come with a stock camera and optionally a LED flash. I used these components to define a system, deriving the heart rate of a person. Figure 2.2 shows the camera setup for measurement.

2.4. Detailed Methodology

The model works on the principle that, every heart beat pertains to a rush of blood in the blood vessels, even in the capillaries at the finger-tips. Whenever the capillaries are rich in blood during a systolic pulse, more light is getting absorbed by the blood, leading to low

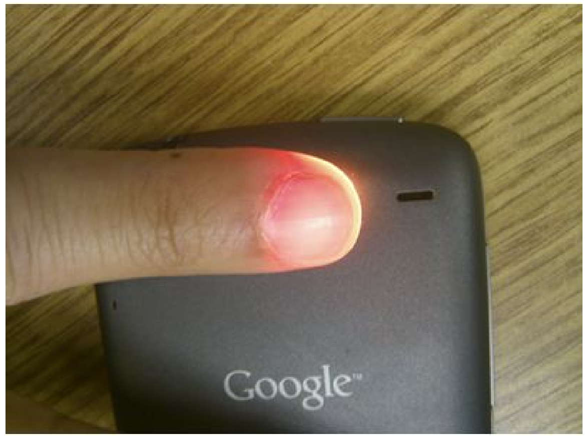


Figure 2.2. *Method of placing the finger over the camera for heart rate measurement*

reflective index and darker frame intensities. Likewise, during a diastolic pulse, most of the light gets reflected leading to bright frames. This change in intensity of light which can pass through the finger creates an alternative pattern of waves similar to a pulse. These changes in intensity with time gives the heart rate of a person.

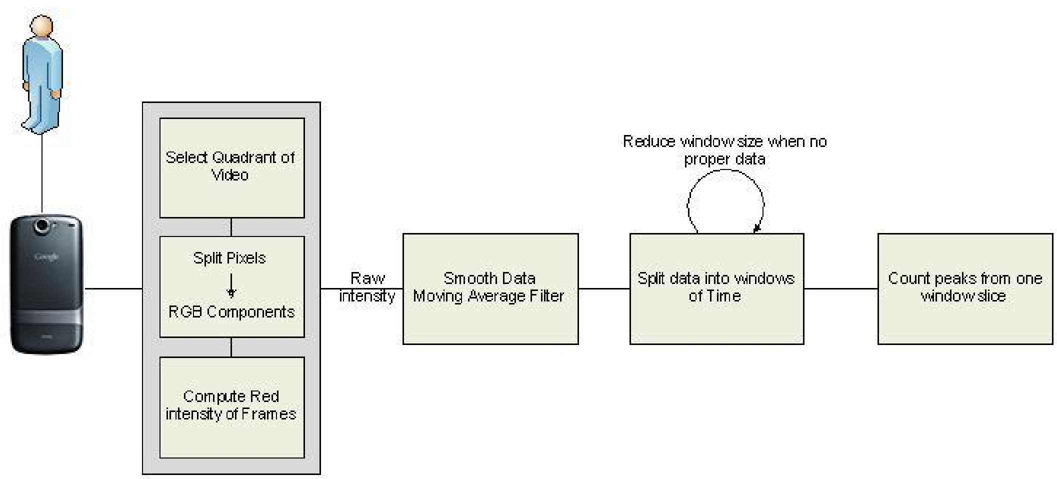


Figure 2.3. *Architecture of heart rate system*

In the proposed method, I record a video of short duration, with the finger placed over the lens of the mobile camera. The flash is turned ON, so that adequate amount of light can reach the finger for proper measurement. For this experiment, I developed an application for Nexus One [5] to keep the LED flash consistently ON while recording video from camera. I discarded the first three seconds of data from the camera, since the CMOS sensor of the camera tries to focus when turned ON. Also, the camera doesn't need to be focused, as the results rely only on the amount of light entering the video feed. It was generally hard to detect the fluctuations in the frames unless the pulses are distinct. A similar methodology was used by Banitsas [6] with a slightly different approach in the analysis of video frames. Figure 5.5 describes the architectural model of the proposed system.

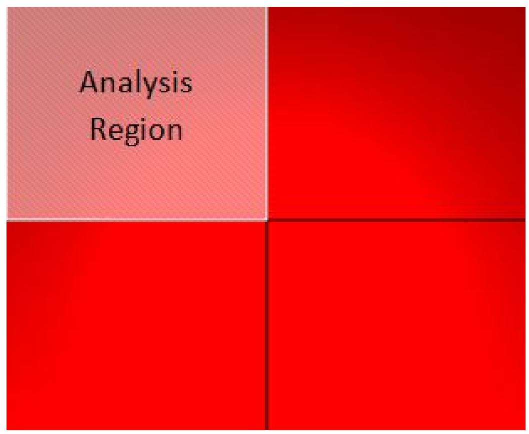


Figure 2.4. *Area under analysis*

2.5. Implementation of Methodology

Working of the system comprises of six functional modules. Initially the video frames were split into four quadrants and only the first quadrant (Figure 2.4) was considered for analysis, since I observed most of the changes and fluctuations are predominant in that region. Every pixel information on each frame was split into individual Red(R), Blue(B) and Green(G) components. In most samples I observed, the prominent color applied only to R with the others tending to zero in every frame, hence difference in the red channel (R_c) intensity to that of all the channels of a frame was negligible. For accuracy of plots, I have considered only the R_c in video frames. The average intensity of pixels for every frame was calculated as its frame intensity. The raw intensity values were filtered with a moving average filter to

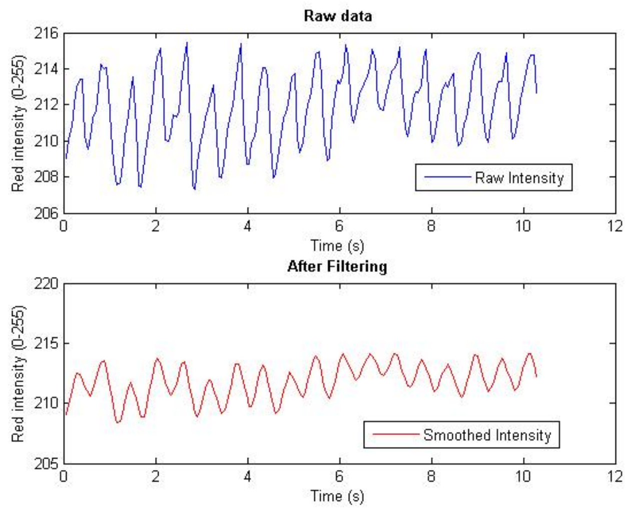
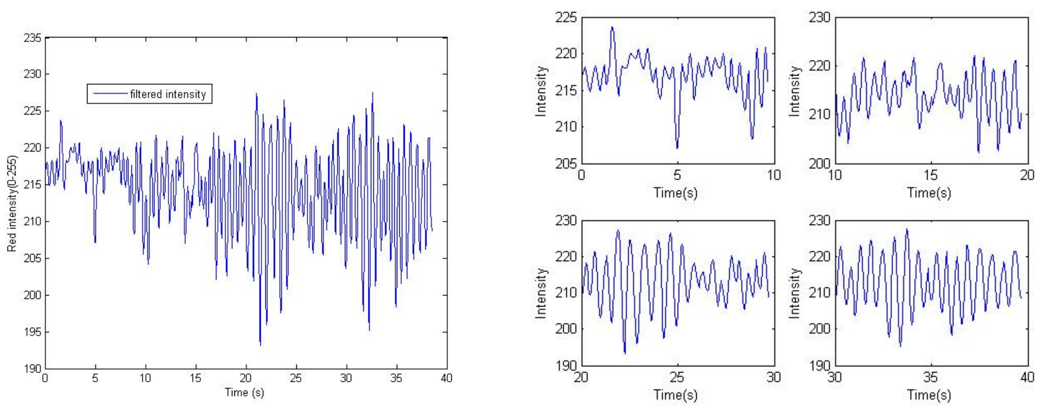


Figure 2.5. *Filtered Data for analysis*

remove rough peaks from the graph for easier identification of peaks. Figure 2.5 shows the filtered results from the raw data obtained from finger pulse. The entire frame was split into



(a) (b)

Figure 2.6. *(a) Total Window of Data (b) Data spilt into smaller Time frames*

windows of fixed length(W_t), for determination of peaks occurring at equal intervals of time as seen in Figure 2.6. If the pattern within W_t matches a sinusoidal pattern, the heart rate is calculated by determining the number of peaks(n) in the window and multiplying the peak count with the window length as given by the equation 1.

$$HR = n * 60/W_t \quad (1)$$

2.6. Heart Rate Accuracy with Proposed Method

This method showed encouraging results with high percentage of accuracy. The collected data was validated with a commercial heart rate monitor available at a fitness center. To prove the effectiveness of the proposed system, I induced higher heart rate to the subject with excessive physical activity. From Table 2.1 it can be seen that the method gives a high

Table 2.1. *Accuracy of results at varying heart rate for a single subject*

	5 sec Window		10 sec Window	
Actual HR	Value	Accuracy	Value	Accuracy
102	108	94.11%	102	100.0%
108	96	88.89%	102	94.44%
114	108	94.74%	114	100.0%
132	132	100.0%	132	100.0%
154	144	93.51%	150	97.40%

percentage of accuracy from the obtained data in finding the HR of the person. It can also be seen that, with the increase in window size for analysis, the error propagation decreases to a very minimum. For better accuracy of the result, the user should hold the finger over the camera lens for a longer time. Table 2.2 shows the expected error in the system when varying the window size. For less error in the results for considered samples, the window size should be kept large. Even more, the accuracy of the heart rate for a full minute of data is precise with the actual heart rate measured manually, with a 100 % accuracy all the time. Based on the need, it is possible to measure every single heart beat with precision.

Table 2.2. *Expected error from the system based on window size*

Window Size (sec)	Expected Error	Error %
5	± 12	$0 - 11.1$
10	± 6	$0 - 5.6$
15	± 4	$0 - 3.7$
20	± 3	$0 - 2.8$
30	± 2	$0 - 1.8$

2.7. Window Time Calibration for Accuracy Improvement

The obtained results look promising, but there is a lot of error introduced due to the smaller window size. If the algorithm misses counting a peak that should be present but moved to the next window in a small fraction of time, the obtained results could vary a lot. Hence I require calibration of the observed results. There are two ways of calibration:

(i) Average

(ii) Window time calibration

Of the two, measuring the average of windows will be the easiest of the calibration methods. Two or more windows could be taken and the average of number of peaks could be computed to give a better result. However, there is a small disadvantage in incorporating this method. If there is not enough legible data available for calculating peaks in multiple windows, the results will be erratic. Hence I consider window time calibration more suitable for this application. Figure 2.7 shows a simulated heart pulse wave. The wave could be different in case of irregular heartbeat and illness.

In this technique, I assume that the heart rate data set is a perfect sinusoidal wave with equal interval between peaks. Based on my observations from the data, I propose algorithm 1. In summary, given the frame intensity value of a fixed time window, the number of peaks

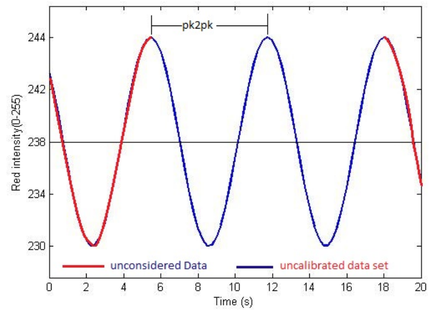


Figure 2.7. *Calibration Of heart rate*

Algorithm 1 Window time Calibration

INPUT: Video Intensity for a Time window (T)

OUTPUT: Calibrated Heart Rate(CHR)

$N \leftarrow no.of peaksinWindow$ /*$N \times \frac{60}{T}$ Gives uncalibrated HR value*/

$pk2pk \leftarrow AvgPeaktoPeaktime$ /*Time between two consecutive peaks*/

$E_x \leftarrow T - N \times pk2pk$ /*E_x is non considered time for Datapoints*/

if $E_x \geq pk2pk$ **then**

$N \Leftarrow N + 1$

end if

if $E_x \geq pk2pk/2$ **then**

$N \Leftarrow N - (pk2pk - E_x) \times 60/T$

end if

$CHR \Leftarrow N \times 60/T$

for the window are counted and the time taken between consecutive peaks is calculated. Then I compute the average value for peak to peak time (pk2pk). If the occurrence of the peaks is harmonic, the average will be same as an individual consecutive peak time. This gives the time taken for one complete cycle of systole and diastole. The wave marked in red in Figure 2.7 is the region that is neglected during the calculation of heart rate. Hence the total time of data which is not being considered is denoted as E_x. When E_x exceeds the time taken for a complete cycle, one more peak is added to the calculated number of peaks and the difference between E_x and the cycle time is taken as the new unconsidered value. If E_x exceeds half the cycle time, I calculate the fractional calibration time value for the peaks, otherwise the number of peaks is taken as the heart rate.

The calibrated heart rate values showed greater accuracy when compared to the uncalibrated data. Commercial heart rate monitors had the same amount of error induced due to time window computation. Hence, the HR for a full 60 sec or 2×30 is considered as the heart rate of a person for comparison with the calibrated results of smaller window sizes. A single user's data set, taken at different time of the day, calibrated for different window sizes is shown in Table 2.3.

From Table 2.3, significant improvement in the results can be seen, much closer to the actual heart rate. Most of the data items had improved accuracy, except a few marked in red in the table which show decreased accuracy from the actual result. This happens due to the irregular heart cycle, where the peak to peak interval varies largely to yield errors in calibration. The longer the data gets, the results will be more accurate. With short data lengths, the chances for error propagation will be high.

2.8. Impact on Heart Rate with Age and Physical Intensity

The approach uses observational data and does not require much computational analysis, however a sensitivity analysis is very helpful in predicting the outcome of the decision based on the situational parameters. In the heart rate measurement, the key parameters involved in the determination of heart are the age and the intensity of work being done. Even though there are some more biological factors involved with it, they cannot be quantified at this point. Researchers from Oakland university [7] have predicted the maximum heart rate of people by age based on records collected over 25 years, giving rise to a non-linear equation

Table 2.3. *Calibration results for Heart rate. HR is the exact heart rate estimated for 60 secs in a dataset; C - Calibrated result; UC - Uncalibrated result; Acc.% is accuracy of Calibrated data; Imp.% is improvement of accuracy in calibrated data from uncalibrated data;*

	5 sec Window				10 sec Window				15 sec Window			
HR	UC	C	Acc.%	Imp.%	UC	C	Acc.%	Imp.%	UC	C	Acc.%	Imp.%
68	72	68	100	5.8	72	70	97.1	2.9	70	70	97.1	0
76	72	82	92.1	-2.6	78	77	98.7	1.3	76	76	100	0
76	72	80	94.7	0	78	77	98.7	1.3	76	75	98.7	-1.3
84	82	86	97.6	0	90	84	100	7.1	84	90	92.59	-7.1
88	82	96	90.9	-4.1	78	82	93.2	4.5	80	86	97.7	6.8
95	96	92	96.8	-2.1	90	96	99	4.2	92	96	99	2.1
97	96	100	96.9	-2.0	90	96	99	6.1	92	96	99	4.1
105	96	108	97.1	5.7	102	102	97.1	0	100	105	100	4.7
105	108	109	96.2	-0.9	102	106	99.1	1.9	104	105	100	0.9
108	96	108	100	11.1	102	108	100	5.5	100	106	98.2	5.5
118	120	118	100	1.6	114	118	100	3.3	128	120	98.3	6.7
120	108	120	100	10	114	120	100	5	116	120	100	3.3
127	132	128	99.2	3.1	132	128	99.2	3.1	120	124	97.6	3.1
130	120	132	98.5	6.1	132	130	100	1.5	128	132	98.5	0

for determining the HR_{max}.

(2) $$HR_{max} = 191.5 - (0.007 \times age^2)$$

Gellish [8], proved the predicted heart rate lies in a tight range between $\pm 2 - 5bpm$ for average individuals. However, the values vary within a wider range for athletes. The actual

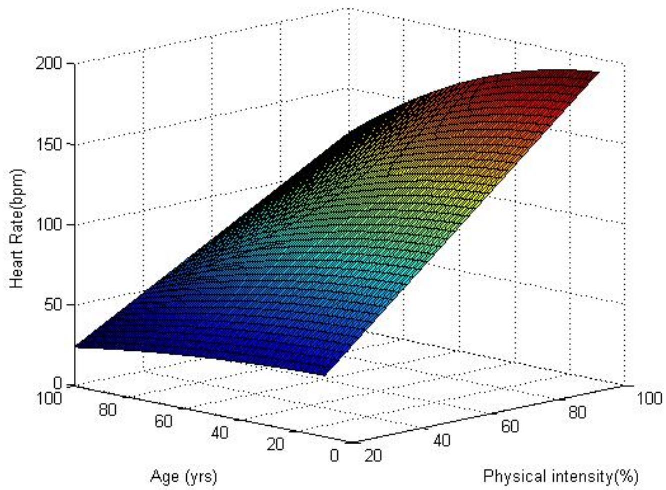


Figure 2.8. *Sensitivity analysis for 10 to 80% increase in Age and Work Intensity*

heart rate depends on the intensity of work involved by the person. Resting heart rate is the lowest heart rate, which can be achieved by a normal healthy person. Even at rest, the intensity of work will be a little more than 25 % and the heart rate increases as the intensity of work increases. There exists a direct relationship between the heart rate and the work intensity, hence the percentage a work intensity for age gives the heart rate at that moment.

From Figure 2.8, it can be seen that the HR gradually decreases with age and increases with work intensity. The resting heart rate of people can go well below normal and reach 45-60 bpm [9] for normal individuals. The results of sensitivity analysis show the normal range of heart rate which a person can achieve based on the age and work percentage. It is also evident that, the proposed methodology follows the pattern and most of the values lie within the range of the graph, proving its effectiveness.

2.9. Conclusion

This chapter I have discussed over the normal medical protocol and described a new system that uses the camera on a mobile phone to find heart rate. From the results, it was inferred that for longer duration of data collection, there was a better chance of achieving more accurate heart rate. It was also observed that the accuracy could be improved from 87% with a 5 sec data to 99 % with a 30 sec data without any calibration. After using the calibration method, it was found that there was an improvement in the accuracy up to 5 or 6 % for even small sized data. This method worked well for 15 sec data and helped achieve 98 % accuracy most of the time.

CHAPTER 3

ESTIMATION OF RESPIRATORY RATE WITH MOBILE SENSORS

3.1. Introduction

Respiration is a physical process which involves exchange of oxygen and carbon dioxide. The main function of the respiratory system is to provide ample oxygen to meet the energy production requirement of the body. A normal ventilation is an automatic, seemingly involuntary action comprising the expansion of the chest cage during inhalation and contraction during exhalation. This normal breathing is relatively constant together with a normal respiratory rhythm. Abnormalities can be detected, when changes are observed in the pattern of rhythm, rate and effort of breathing.

3.2. Medical Protocol

The usual practice for measuring respiratory rate, involves a watch with a second hand or a stop watch. The number of times, the chest moves upward is counted when the person is in a seated position or lying flat for a full 60 seconds or count for 30 seconds and multiplied by two. It is considered optimal to place the hand on the upper chest to feel it rise and fall. Each rise and fall counts as one cycle of respiration. Also while calculating the respiratory rate, a note has to be made about the effort in breathing by the person and the type of breathing.

3.3. Brief Description of Proposed Method

In this section, I propose a new method of measuring the respiratory rate with the mobile phone's built-in accelerometer. The accelerometer is a device to detect small changes in acceleration of the device. The Google's Nexus One uses a 3 axis accelerometer to detect acceleration changes in all the three dimensions. I have observed that the phone is sensitive to detect breathing changes caused by normal effortless breathing. In this method, I place the mobile phone over the person's chest lying flat to the ground for recording acceleration changes due to the movements of chest wall. The acceleration values are plotted to give the respiratory pattern of the person. From the acceleration graph, I was able to detect the type of breathing and the time taken by the person to complete a breathing cycle. Figure 3.1 shows the setup for placing the phone.

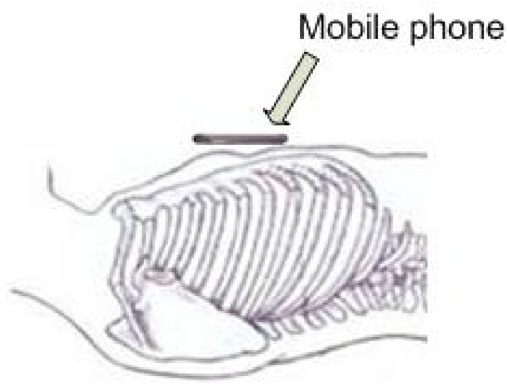


Figure 3.1. *Setup for Respiratory rate*

3.4. Detailed Methodology

The accelerometer data from a human subject lying flat on the ground with his back to the floor was collected to analyze the breathing pattern. The phone was placed in the lower part of the sternum under three different scenarios. When the subject is not breathing at all; when the subject is breathing normally and the subject is breathing heavily and complete. In this setup, the breathing was voluntary and the subject was conscious of the type of breathing. This result would vary from a subconscious breathing to a small degree. Figure 3.2 describes the model of the system for measuring the respiratory rate.

The Nexus one gives an average of 20 - 30 values per second for the accelerometer sensor. Although the sampling intervals were not constant due to internal hardware capability

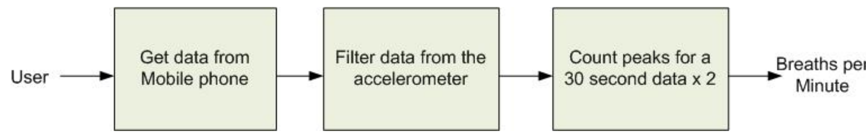


Figure 3.2. *Architecture of respiratory rate system*

issues, the average sampling time was considered for analysis. The BlackBox V2 application records in three different modes or speeds of reporting data, starting from the normal rate to the fastest. It is preferred to set the blackbox mode of operation to fastest, while taking respiratory rate measurements. The application under normal mode of operation did not show significant changes in the acceleration value for shallow breathing. Figure 3.3 shows the

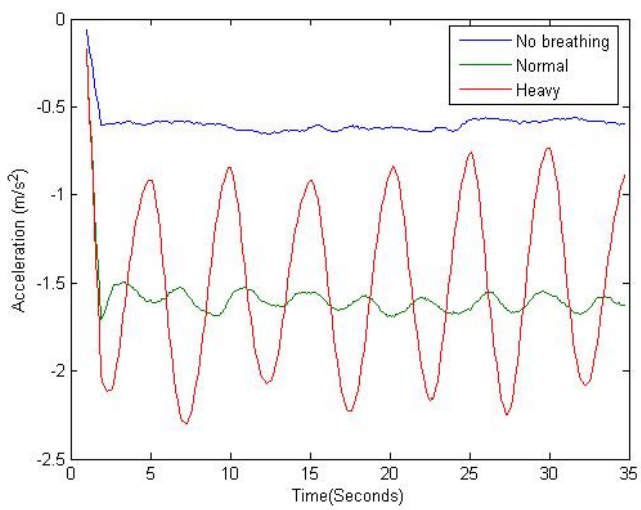
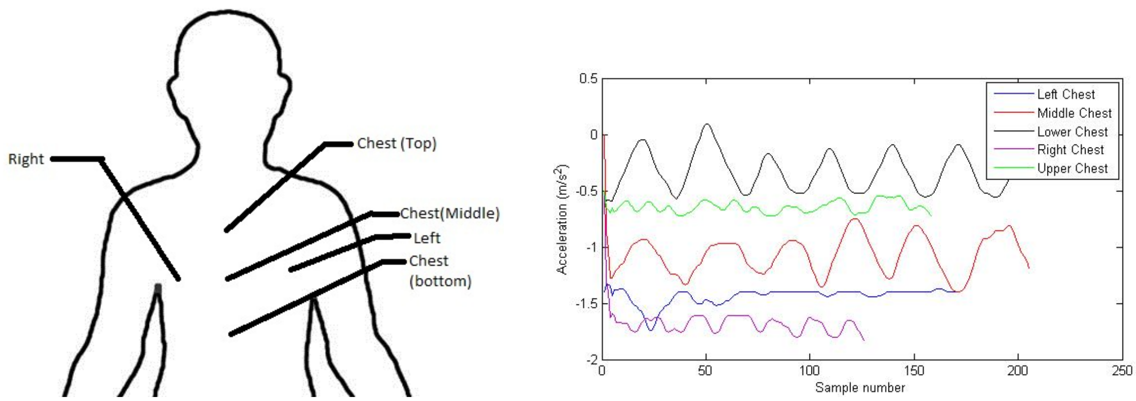


Figure 3.3. *Acceleration data for breathing under 3 scenarios*

different breathing patterns measured by accelerometer reading of the mobile phone. When the user was not breathing there was no significant change in the acceleration except a few disturbances or noise. After careful filtering of the data, I was able to detect the acceleration changes of the subject breathing normally. The amplitude of accelerations was low because of incomplete breathing pattern in terms of lung capacity and air intake. When the subject was breathing heavily, I have observed an alternative pattern of acceleration varying within a wide range. Heavy breathing had large acceleration changes when compared to normal, but the frequency of breathing was low since it took a longer time to inhale and exhale fully. The pattern projects a harmonic sinusoidal wave, representing the pattern of breathing.



(a) (b)

Figure 3.4. *(a) Different regions selected over the chest to determine the place with maximum impact (b) Smoothed plots for the acceleration values on different regions when breathing normally*

Figure 3.4 (a) shows the setup which is carried over on different places on the chest to find the location giving the most significant results. The mobile was placed over the lower chest, middle chest, upper chest, right chest and left chest as shown in the figure. The results obtained from Figure 3.4 (b), show that the maximum distinction of results were from the middle and lower part of the chest, which is considered ideal for placing the mobile while measuring.

Figure 3.5 shows the difference in pattern between hyper ventilation and normal breathing, which are clearly being distinguished. The values are taken along the X- Axis even though there was a little effect of acceleration on all three axes. Hyper ventilation is the stage when the person is not able to breathe properly and regularly, characterized by short and shallow respiration.

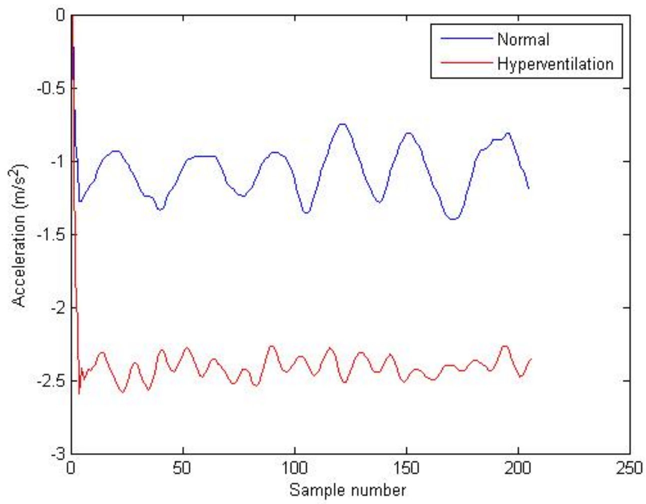
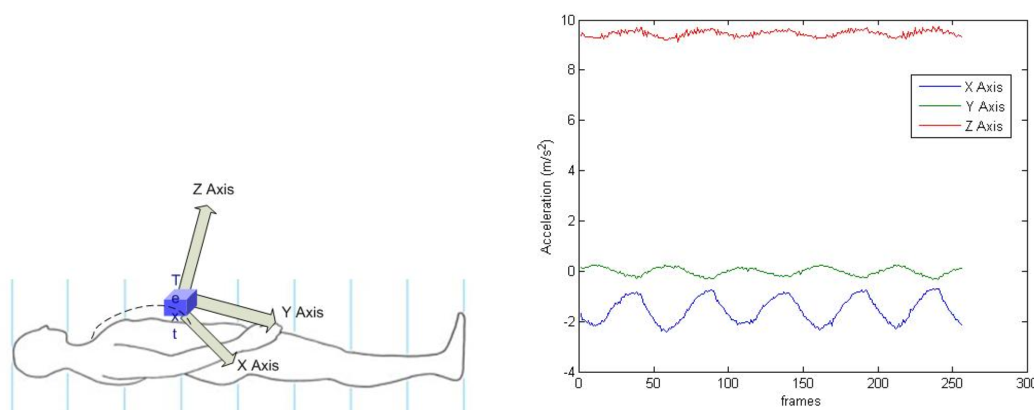


Figure 3.5. *Comparison of normal rate with hyper ventilation*

The orientation of the phone determines the values for the accelerometer. The human body is not exactly flat to measure the acceleration signal in the Z axis alone. When a person inhales, the lungs create a curvilinear motion causing it to expand. While breathing, the body creates movement in at least two axes or all three axes which can be seen clearly from Figure 3.6 (b). The displacement on the X axis was more than the Z axis because



(a) (b)

Figure 3.6. *(a) Displacement of the chest cage in circular fashion on inhale(b) Plot showing the acceleration changes on all the axis*

during an inhale process, the mobile phone moves towards the direction of the neck than the upward movement introduced in the Z axis(Figure 3.6 (b)). This displacement is often unnoticed and considered as up and down motion. The values will differ from person to person, but the respiratory rate and the pattern generated by the inhalation and exhalation will be similar for every healthy person. There are some respiratory disorders which can be diagnosed by the pattern of the respiratory action. Often, these patterns are hard to detect by manual assessment (Figure 3.7 [10]). My proposed method will be best suited for diagnosing some of the respiratory disorders. With simultaneous differential measurement of breathing patterns from two mobile phones in the chest and abdomen, it is possible to detect respiratory disorders during an emergency or accident. Similarly, abnormal breathing, suffocation, etc., can be detected from the breathing pattern with the proposed method.

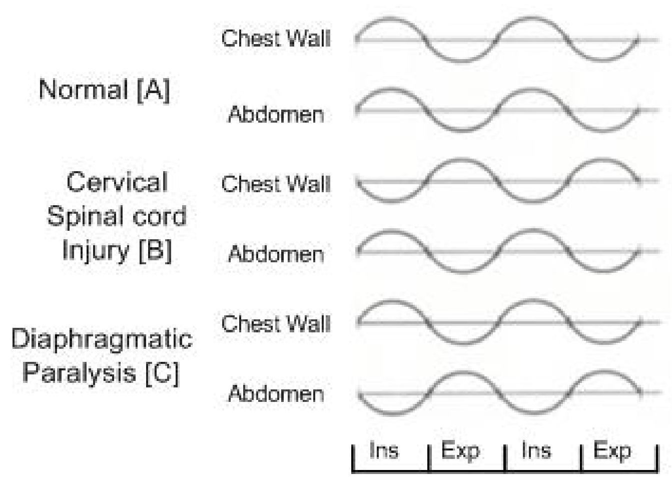


Figure 3.7. *Respiratory disorder diagnosis*

3.5. Respiratory Pattern from Oxygen Saturation

Though the respiratory rate can be detected using the accelerometer sensor, its is difficult to make the subject lie flat to the ground for checking the respiratory pattern. During the tests conducted for the heart rate analysis (Chapter 2), the results showed interesting pattern of waves which closely reflected the respiratory pattern of the subject. I have done a detailed investigation on the occurrence of this pattern and science behind it. The pulse recorded by the mobile camera closely resembles the pulse oximetry [11] in working and results. On other hand, pulse oximetry uses the same model followed by the proposed heart rate determination. Pulse Oximetry is a simple method of determining the percentage of haemoglobin in the blood that is saturated with oxygen. The oximeter has probes connected to the finger which can display the percentage of oxygen in the blood for heart beat. The probes contain two LEDs which emit light of known wavelength in red spectrum and infra-red spectrum. When the lights enter the finger, some amount of light is absorbed by the blood and tissues depending on the concentration of haemoglobin from which the pulse and oxygen level are determined.

In my proposed method for heart rate determination, I made the mobile camera to act as the photo receiver and the LED flash to resemble the LEDs of oximeter probe. Currently, my method can detect the oxygen saturation level in blood by observing pulse from the finger-tips, but it cannot be completely relied upon for the calculation of percentage of oxygen in blood. When an individual inhales, the heart is supplied with a large amount of oxygen to be delivered to the tissues for metabolic activity, hence the amount of haemoglobin saturated with oxygen increases. The count of saturated haemoglobin decreases, when the person exhales. These changes in oxygen saturation levels were clearly distinguishable from the results of heart rate.

The methodologies and working of the method were discussed in Chapter 2, hence the following section discusses only the results and patterns obtained from the heart rate results.

3.5.1. *Discussion on Respiratory Patterns*

The derivation of breathing pattern requires filtering of pulse wave. In the derivation of respiratory pattern, I pass the camera feed from the subject's finger through a 10^{th} order running median averaging filter implemented as a smoothing function in MATLAB. The resulting output shows the respiratory pattern of the individual. Figure 3.8 shows the working of respiratory pattern detection from finger pulse. Figure 3.9 shows the respiratory pattern obtained from a subject counting the number of air intake and outflow. The pattern showed the number of intakes and outflows precisely and I have also observed from the figure, when the subject was breathing hard and shallow.

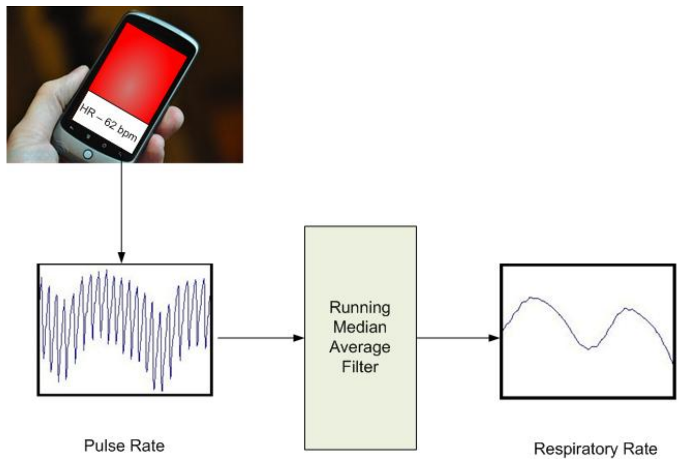


Figure 3.8. *Model of respiratory pattern from Oxygen Saturation with mobile phone*

Based on results obtained from subjects at different time, it was also observed that the respiratory pattern obtained for individuals varied based on the type of breathing and it was hard to detect slow and shallow breathing. Leonard et. al [12] proved the usage of standard pulse oximetry permits accurate determination of respiratory rate among health adults. Arnold et. al [13] estimated the severity of airway obstruction using plethysmography in patients with obstructive airway diseases. Hence it should be possible for the proposed methodology to detect the respiratory rate and substitute photo plethysmography.

3.6. Conclusion

In this chapter, I have discussed some standard methods of checking respiratory rate in medical and emergency conditions, along with the introduction of a new system for using mobile phone's accelerometer to estimate respiratory rate. It was inferred that, its use can help in accurate determination of the respiratory rate. Also I have demonstrated the usefulness of the accelerometer in detecting abnormalities that cannot be observed by the traditional

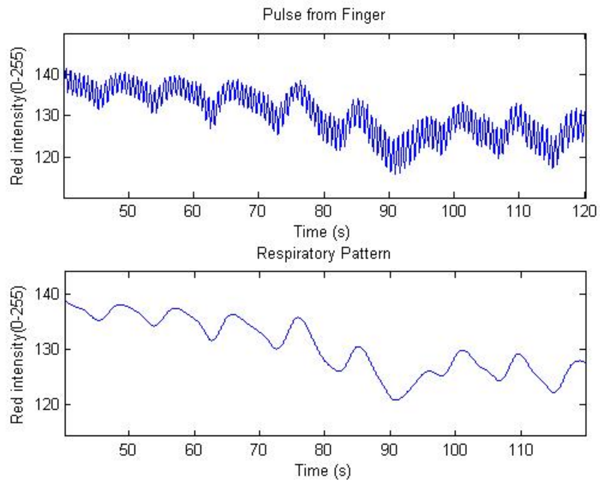


Figure 3.9. *Derivation of Respiratory pattern from finger pulse*

methods. However, I am planning to make the methodology more robust and develop new techniques for identifying oxygen saturation in blood using the mobile phone. In future, this application can potentially replace the traditional pulse oximetry.

CHAPTER 4

CUFFLESS BLOOD PRESSURE ESTIMATION WITH MOBILE PHONE

4.1. Introduction

Measurement of vital signs or parameters in a human is an arduous task when sudden dizziness or fainting could occur during unexpected situations. The only information that we are aware of is that these are the most common symptoms of low blood pressure. However, to check a person's blood pressure in such situations, we need for a portable, convenient device or apparatus. Despite the availability of digital wrist and arm blood-pressure meters, most people do not carry these devices during their daily travel to work place, gymnasiums, recreating facilities, etc.

Blood pressure, the amount of force applied on the walls of the arteries when the blood is forced throughout the body,depends on factors such as the amount of blood in the body, the pumping rate of the heart, the flexibility of the arterial walls, and resistance to blood flow due to the size of the arteries. A continuous variation of blood pressure occurs in a human due to factors such as physical activity, medication, anxiety, and emotions. The body has unique mechanisms to regulate a person's blood flow; whenever a person's blood pressure drops, the heart rate increases to pump more blood and the arterial walls contract to increase the blood pressure. The blood pressure is given by two numbers measured in millimeters of mercury (mmHg).

- The systolic pressure represents the amount of pressure applied in the arteries when the heart contracts.
- The diastolic pressure represents pressure on the arteries when the heart is at rest.

Most non-invasive methods give an approximate value of pressure for an instance and it's accuracy depends on the ability of the physician to determine its value.

4.2. Terminologies Used

- SV - Stroke Volume: The Volume of blood pumped from the heart in one beat.
- P_m - Mean Arterial pressure: Medical term to describe the average blood pressure of an individual.
- P_s - Systolic Pressure: The blood pressure when the heart is contracting.
- P_d - Diastolic Pressure: The blood pressure when the heart is relaxing.
- P_p - Pulse Pressure: The difference between the systolic and diastolic pressure.
- HR - Heart Rate/Pulse rate: The number of heartbeats per minute.
- Z - Impedance to blood flow: The total opposition of blood in the vascular system for blood flow.
- R - Resistance to blood flow: The measure of blood opposition in the blood vessel for steady flow.
- N_a - Difference in video intensity within a single systole and diastole.
- S_a - Peak amplitude in video intensity when the cuff pressure equals systolic pressure.
- D_a - Peak amplitude in video intensity when the cuff pressure equals diastolic pressure.
- ET - Ejection Time: Time taken for the opening and closure of the aortic valve.
- BSA - Body Surface Area: Physiological term to measure or calculate the surface area of the human body.
- Q - Cardiac Output: The volume of blood being pumped by the heart in a time interval of one minute.
- VTT - Vascular Transit Time: The transmission delay of the first heart sound to be felt as a pulse in finger.

4.3. Traditional Protocols

Every one should have their blood pressure checked regularly. Though blood pressure is normally taken in the doctor's office during each visit, care has to be taken for certain groups such as people over 35 years of age, people with history of high blood pressure, pregnant women, etc. Hospitals and medical response teams follow certain techniques and procedures for blood pressure measurements. This section provides a brief discussion on those medical protocols.

Blood pressure is typically measured in the upper arm with the person comfortably seated and the arm in level with the heart. The measurement apparatus is a mercury sphygmomanometer [14] comprising a manometer, pressure cuff, bladder and a gauge to show the pressure value inside the cuff. The cuff is wrapped tightly around the upper arm about an inch from the elbow since the upper arm is closest to the heart and errors due to upflow/downflow of blood are eliminated.

The physician locates the brachial artery in the elbow and places a stethoscope over it to listen for sounds produced during the process. When this setup is done, the cuff is inflated by pumping air through the bladder to a pressure of about 210 mmHg or usually 20 to 30 mmHg above the normal systolic pressure. At this point, the physician will not hear any sounds produced by the blood flow. The physician gradually reduces the pressure in the cuff at a rate of 2-3 mmHg/second. At a certain threshold, the artery opens slightly, counteracting the pressure induced by the external cuff. This action is followed by sounds due to the turbulent flow of blood. The cuff pressure corresponding to the first sound produced is taken as the systolic pressure. The cuff pressure is further reduced and when the sounds due to blood flow cannot be heard with the stethoscope, the physician takes that cuff pressure as the diastolic pressure. This procedure may be repeated twice or thrice in the same day to determine an accurate value. Figure 4.1 shows the blood pressure apparatus setup during a measurement process.

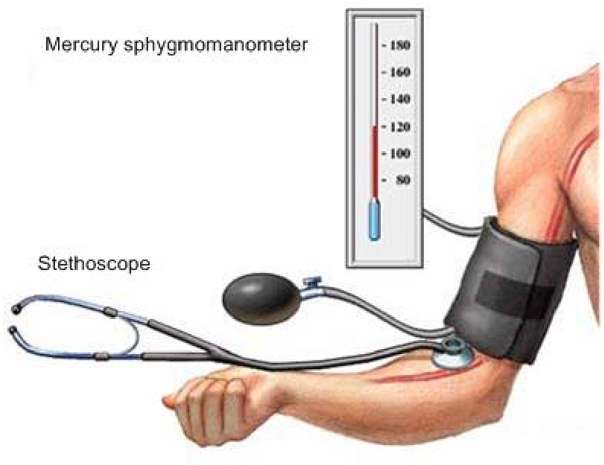


Figure 4.1. *Traditional setup for estimation of blood pressure by sphygmomanometry (img src:Adam Inc.)*

Other non-invasive methods like the ECG and Pulse Oximetry can give a continuous blood pressure value more accurately at any instance of time. This capability proves to be useful during surgery.

4.4. White-coat Syndrome

White-coat syndrome [15] is a phenomenon where the blood pressure is always at a higher value from the normal. A patient experiences this due to nervousness and over-reaction to the environment. It also occurs due to repetitive inflation of the cuff consecutively over the arm. The repetitions induce stress in the person; resulting in a higher or lower blood pressure than the expected normal value.

4.5. Brief Description of Proposed Methods

Two methods are proposed for determining blood pressure using mobile phones and their built-in sensors. The first method is a cuff-based blood pressure assessment which is closely related to the traditional way of measuring the blood pressure, and the second is a cuffless estimation using the mobile phone's built-in audio and video sensors. Care has been taken to select test subjects presenting no medical history of cardiac disorder. A detailed discussion about the methods of blood pressure estimation is provided in the next section.

4.6. Cuff Based Estimation

This method resembles the traditional method of measuring blood pressure, but replacing the stethoscope with the mobile phone's camera sensor. In addition to the video data, an automatic wrist blood pressure meter's cuffing system [16] is used, delivering the time for pulse arrivals. For estimation, the arrival of pulse signal from the video data is checked. The amplitude of the peaks is of significant use in this method. The blood pressure of a person is determined by evaluating many physiological factors such as stroke volume, heart rate, resistance of the blood to the flow, and viscosity.

4.6.1. *Detailed Methodology*

This estimation method follows a pattern similar to the auscultation techniques(described in Section 4.3) to listen for the internal sounds of the body, where a camera sensor is used to view the blood flow in the arteries. Figure 4.2 shows the test setup for this method of estimation, where the automatic blood-pressure meter's cuff produces auscultation effects in the wrist which are then reflected in pulse waves in the finger tips. The determination of blood pressure follows a new formulation involving factors contributing to the equation. The systolic pressure depends on the mean arterial pressure and the person's pulse pressure. Therefore, as a first step, the pulse pressure must be determined.The pulse pressure is computed by correlating the time difference between the occurrence of systolic and diastolic pulse in the graph. I formulated a regression equation based on the pressure readings on the meter to the time it takes to reach the value. Figure 4.3 (a) shows the curve fit to the data for a decrease in cuff pressure on the automatic cuff system. By applying the regression equation,

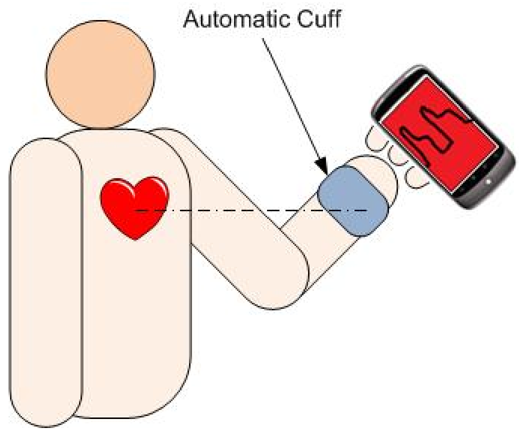
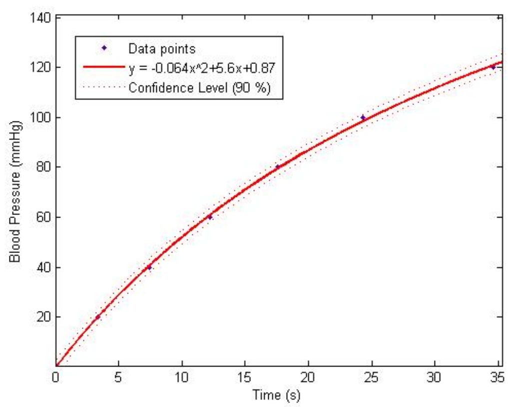
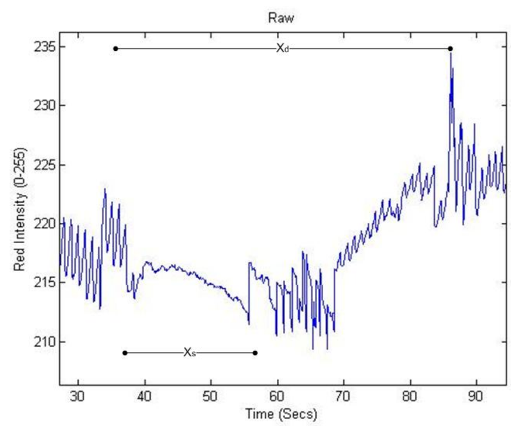


Figure 4.2. *Test setup for cuff-based estimation: The subject is in the sitting posture with the heart inline to the pressure cuff. Pulses are measured over the finger with mobile camera*



(a)



(b)

Figure 4.3. *(a)Regression relation for the pressure value on the meter, corresponding to the time it takes to reach that value from full inflation of cuff (b) Determination of* X_s,X_d *which are the time taken to reach systolic and diastolic pressure on the cuff respectively.* y_s *and* y_d *were determined from* X_s,X_d *in 3(b) based on the regression equation derived in 3(a)*

I can determine the blood pressure at a certain instance of time.

(3) $$y = -0.064x^2 + 5.6x + 0.87$$

The proposed approach utilizes only a mobile phone and an automatic cuffing system; hence, the pulse pressure has to be determined using a naive approximation method. The results obtained through the phone along with the cuff mechanism clearly show the distinctive regions of the systolic and diastolic phase consistently; therefore, this information from the graph can be used to predict the pulse pressure. The pulse pressure is obtained by taking the change in pressure between systolic and diastolic time. Therefore, the pulse pressure is calculated by

$$P_p = y_d - y_s \quad (4)$$

where y_d and y_s are calculated based on the regression equation using x_d and x_s from the Figure 4.3. From the systolic and diastolic peaks observed, the pulse pressure values were calculated using equation 4. The regression equation changes for differing cuffs and will be linear if the cuff pressure reduced steadily and gradually.

Even though the measured values of pulse pressure are approximate, the estimated values reveal closeness to actual values for some datasets. The person's heart rate was determined by taking the video sample of around 15 - 30 seconds before inflating the cuff. Now having obtained the pulse pressure, I can calculate the stroke volume from equation 8 by applying the person's corresponding age and weight. When stroke volume is considered as a constant over a period of time, the pulse pressure must stay constant since stroke volume and pulse pressure are directly related.

Table 4.1 shows the accuracy of finding the pulse pressure with the proposed method, conducted on five subjects at different times. I was able to observe that, the calculated accuracy of pulse pressure varied between 50 - 100 % over the datasets.

Having the pulse pressure, I can calculate the mean arterial pressure for the subject using the equation

$$P_m = HR \times SV \times Z \quad (5)$$

Table 4.1. *Estimated Accuracy of Pulse pressure for five different subjects*

Subject	Measured P_p	Actual P_p	Accuracy %
Sample 1	30	20	50
Sample 2	22	21	95.2
Sample 3	30	24	75
Sample 4	36	36	100
Sample 5	44	48	91.67

Impedance [17] is a major factor in the case of impedance cardiography technique for determining a person's blood pressure.

(6) $$Z = \frac{R}{1 + i\omega RC}$$

Due to unavailability of impedance graphs, I created a crude relation for describing the resistance of blood flow with the amplitude of the waves.

Now, the impedance factor Z has to be determined to calculate the mean arterial pressure. Because Z is calculated only for impedance cardiograph, an alternative method was required to determine the impedance factor for a dataset. In my proposed system, data is normalized and a crude value of Z is calculated using

(7) $$Z = \frac{N_a}{S_a - D_a}$$

where S_a, D_a denote peak amplitude in the video frame at the systolic and diastolic time instants respectively, and N_a represents the difference in video intensity during one systolic and diastolic cycle. The values of N_a, S_a, and D_a are obtained from the Figure 4.4

Calculation of Z is based on an assumption that, the ratio between a normal blood flow without external resistance to the difference between the peak amplitudes with the maximum

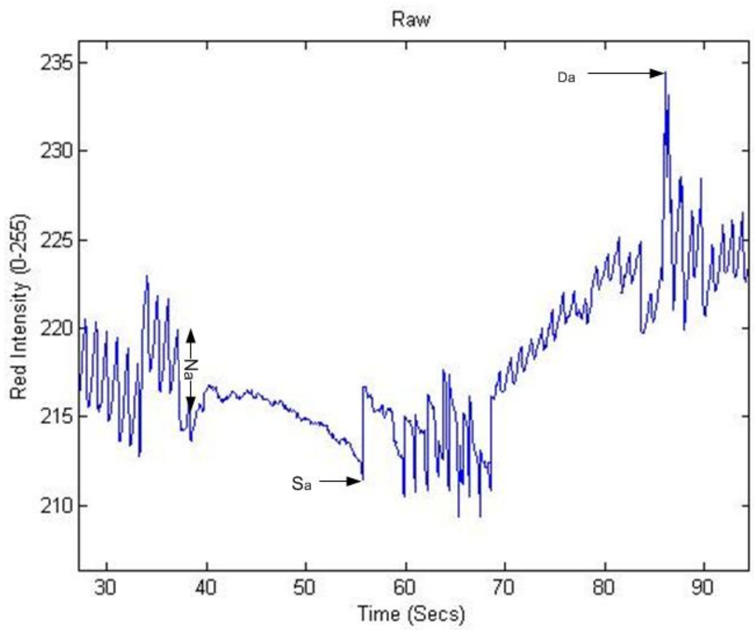


Figure 4.4. *Estimation of Z. Values of N_a,S_a,D_a are obtained as shown and substituted in equation 7*

external resistance gives the impedance factor for the flow. However, currently, no test equipments exist to support my assumption of the impedance factor theory.

A person's stroke volume (SV) is determined based on the approximate amount of blood in the person and their pulse pressure. Equation 8 [18] gives the value of stroke volume(the blood pumped by heart in each beat).

(8) $SV(mL) = (0.013 \times Bodyweight(kgs) - 0.007 \times Age(years) - 0.004 \times HR + 1.307) \times P_p$

With the help of the mean arterial pressure and pulse pressure, the systolic and diastolic pressure can be determined. In medical terms, pulse pressure is determined from the impedance cardiogram during a continuous pressure measurement.

$$P_s = P_m + \frac{2}{3} P_p \tag{9}$$

$$P_d = P_m - \frac{P_p}{3} \tag{10}$$

Where P_m and P_p are mean arterial and pulse pressure, respectively.

Having values for the necessary variables, I can estimate the systolic and diastolic pressure by substituting these values in the equation 9 and 10. Based on my assumptions and analysis, I determined the blood pressure values (tabulated in Table 4.2).

Table 4.2. *Accuracy of P_s and P_d for five different individuals*

	Measured		Actual		
Subject	P_s(mmHg)	P_d(mmHg)	P_s(mmHg)	P_d(mmHg)	Mean Accuracy %
1	90	60	85	65	91
2	76	54	87	66	84.5
3	74	44	89	65	75.5
4	106	70	102	66	96.1
5	130	86	123	75	89.7

From Table 4.2 it can be seen that the pressure values are imprecise, but a person's bp level (high, medium, low) can still be determined. However, the magnitude of error increases rapidly, when the determined values of the variables are erroneous. I validated my measurements with the help of a commercial blood pressure meter, but this cannot guarantee a high level of accuracy. All the tests were done multiple times to determine the consistency of the plots.

4.7. Cuffless Estimation

During an emergency, it is not possible for users to carry special devices for diagnosis. The use of a cuff(described in Section 4.6) in the blood pressure estimation voids the purpose of this application. Hence, I propose an alternative method in this section, for determining a person's blood pressure without the use of a cuff.

4.7.1. *Detailed Methodology*

The measurement methodology for cuffless blood pressure estimation resembles the heart-rate estimation discussed in Chapter 2; however, additional care must be taken while holding the finger over the camera. The camera's lens should be placed exactly in the center of the finger with a moderate amount of pressure applied over the lens. If the finger is held too tightly against the lens, amplitude variations in the video will be very small; likewise, when the finger is kept loose, the video tends to be dark all the time. For measurement simplicity, I had a single subject under-went multiple tests to verify the measured data's legitimacy.

Initially, the person's stroke volume must be determined to compute pulse pressure. Stroke volume is given by the equation 11 [19]

$$SV(mL) = -6.6 + 0.25 \times (ET - 35) - 0.62 \times HR + 40.4 \times BSA - 0.51 \times Age \tag{11}$$

where ET(ms) is the ejection time, and BSA (Body Surface Area) is given by

$$BSA = 0.007184 \times Weight^{0.425} \times Height^{0.725}$$

Ejection time is the time difference between the opening and closure of the aortic valve during a single pulse. This ejection time is estimated based on information from a graph obtained by analyzing the video. The intensity of the video frame increases when there is no blood flow; hence, the dips in the graph represent actual blood flow identified from the video frames. To determine a systolic and diastolic peak, I invert the amplitude of the obtained graph for easier analysis. Figure 4.5 shows the determination of the ET value from the graph for a single pulse recorded from a video stream.

Nürnberger *et al.* [20] in their paper on determining pulse wave velocity from Left Ventricular ejection time have used the phonocardiogram(electronic stethoscope) for estimating

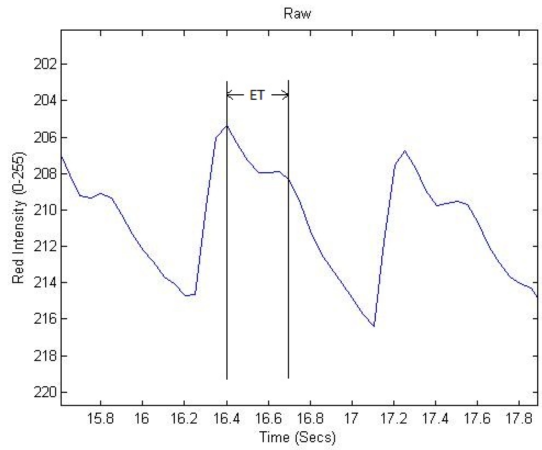


Figure 4.5. *Video intensity of a single pulse: Peaks corresponds to opening and closure of aortic value in the heart. The difference in time between opening and closure gives Ejection Time (ET)*

ET among 102 subjects. They have explained in their architecture, the methodology for observing audio data to derive ET. In my proposed setup, I make use their method for ET estimation using a stethoscope. I present a correlation between the stethoscope data and video data collected over the phone to provide the localization of systolic and diastolic phase in a pulse obtained from video data. Figure 4.6 shows the region corresponding to the opening and closure of the aortic valve. The time difference between the two gives the ejection time.

With the data obtained from the subject, I determine the person's stroke volume. Because stroke volume and pulse pressure are directly related, the pulse pressure was calculated using the equation 12(specified in [21]) and represented in units of mmHg.

$$P_p = \frac{SV}{(0.013 \times Weight - 0.007 \times age - 0.004 \times HR) + 1.307} \tag{12}$$

Next, the mean arterial pressure P_m of the person was determined based on cardiac output (Q) and resistance to blood flow(R). Cardiac output is the amount of blood flowing through the heart for one minute usually measured in L/min with the average male cardiac output

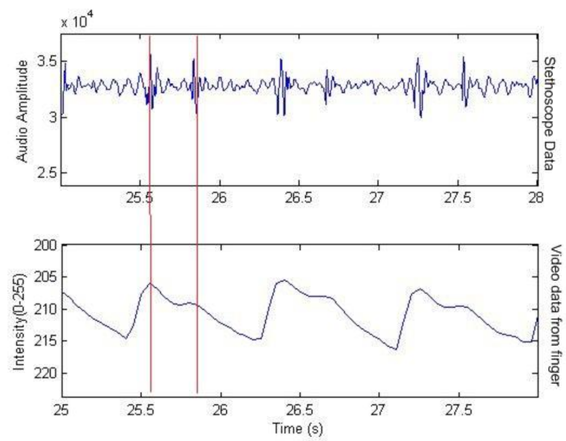


Figure 4.6. *Correlation of Stethoscope data with video intensity from camera measuring ET*

calculated to be 5 L/min and the average female calculated to be 4.5 L/min [22] where L is in liters.

$$P_m = Q * R \tag{13}$$

The flow of blood is inversely proportional to the resistance of blood. Resistance, a complex factor, depends on the length of the vessel, vessel's size/radius, and viscosity of blood. The relationship between the blood flow and resistance is given in terms of fluid mechanics by Poiseuille's Law [22] given as

$$Q = \frac{\Delta P r^4 \Pi}{8 \eta l} \tag{14}$$

where ΔP is change in pressure , r is the vessel radius, η is the blood viscosity, and l is the vessel length. From the proposed system architecture, it is not possible to use the traditional equation to determine person's cardiac output. Hence, I define a system, that can present the range of the person's pressure value by using a reference data set.

Figure 4.7 shows my method for determining the value of resistance from a known data set, where pressure values were measured using a traditional pressure estimation. The resistance of blood to the flow for a particular instance cannot change within a short time span. Thus,

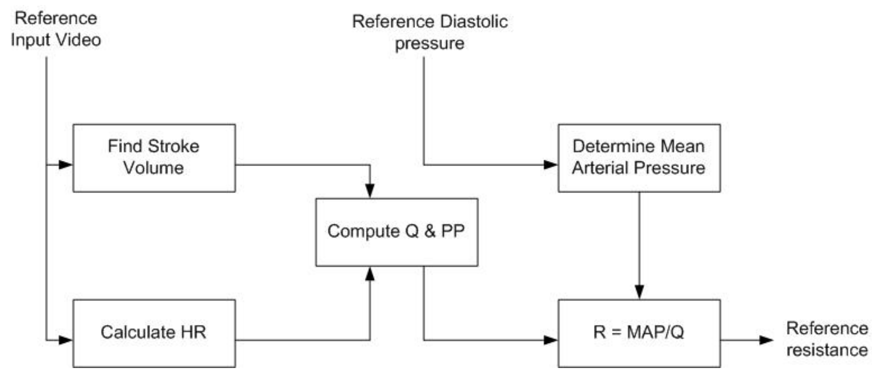


Figure 4.7. *Methodology for estimation of reference resistance from a known blood pressure value of subject*

I calculated a person's resistance with a known pressure value taken at complete rest and fixed them as a constant for all other readings of that subject.

A person's blood pressure cannot stay constant throughout the day, and the value of resistance changes as the body strives to maintain the blood pressure at a normal value. The results obtained from my analysis and observations from the resulting data are discussed in the following section.

4.7.2. *Accuracy with Cuffless Estimation*

I have performed a series of experiments with multiple subjects; however, for discussion purposes, the following results reflect the values and accuracy derived from a single subject. The individual's resistance to blood flow was determined based on the method followed in Section 4.7.1. Table 4.3 presents the data used in the calculation of reference resistance value of a single subject.

Using the parameters tabulated in Table 4.3, the resistance to blood flow at a particular instance was determined as 18.31 mmHgmin/L. By assuming the resistance value as a constant for the other datasets, I could then calculate the blood pressure value from the ejection time(which was in turn measured from the video intensities). Table 4.4 shows the results obtained from one subject. The measurements were taken at 30 min intervals and the results

Table 4.3. *Test Subject information*

Weight	140 lbs	Height	5.8 ft
Heart Rate	75 bpm	Ejection Time	234 ms
Systolic Pressure	95mmHg	Diastolic Pressure	64mmHg
Age	23 yrs	Pulse Pressure	31 mmHg

were derived using a constant resistance value. This method, it should be noted, can give rise to erroneous results since the resistance of blood cannot stay constant for a long time. Blood pressure due to normal conditions will be compensated for by an alteration of the flow resistance.

Table 4.4. *Accuracy of Pressure estimated based on a constant resistance to blood flow R= 18.31 mmHgmin/L, Measured represents the values estimated and Actual represents the value obtained using a blood pressure meter*

HR	Measured P_s	Measured P_d	Actual P_s	Actual P_d	P_s error %	P_d error %
68	100	64	99	66	1.2	2.8
69	99	63	107	78	7	18
66	98	62	104	73	5	14
62	97	60	90	63	8.5	4.1
60	119	72	90	60	32.6	21.29
60	78	47	89	64	11	25
75	99	65	99	63	0	4
74	125	82	93	66	35	25
75	101	66	98	64	3.3	4
72	77	50	90	65	13	21

From the results, I was able to obtain an improved accuracy in the measured blood pressure value from the Cuff-based pressure estimation discussed in section 4.6. Even though the cuffless blood-pressure estimation method gave an accuracy between 65-100%, the accuracy can not hold for over a period of time with the same resistance values. In reality, a person's blood flow resistance will change due to anatomical and psychological variations over time. In my test setup, the measurements depicted in the table were taken on a single day; and, the accuracy variations caused in the results might be due to the change in resistance value to the blood at that instance. Since the resistance was assumed as a constant throughout, the estimated pressure varied widely. It was also observed that pulse pressure in the measured data was close to the subject's actual pulse pressure and had an error of 10% to the maximum.

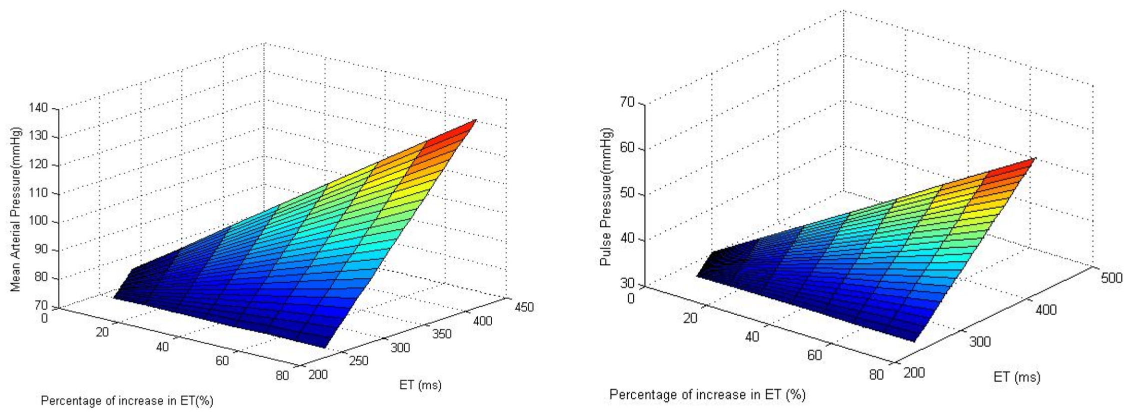
Test results showed similar behavior for other subjects. Though the working of the methodology can be accepted to a certain extent, the accuracy and randomness of the results due to physiological changes in the vascular system defy the usability of cuffless estimation system on all occasions. While this method can provide a subject's pressure with a reference resistance value and it will be suitable for defining the range of pressure. To improve the accuracy of the results, I plan to use differential measurements of heart sounds and pulse waves, as discussed in Section 4.8.

4.7.3. *Sensitivity Analysis for Cuffless Estimation Method*

The results obtained using cuffless estimation of blood pressure showed some closeness to the subject's actual blood pressure. A detailed analysis on the procedure's sensitivity can reveal information about the outcomes and situations that encourage errors. The cuffless estimation depends completely on the proper identification of systolic and diastolic peaks from a pulse wave with video intensity. When the time taken for the systolic and diastolic peaks is misreported, resulting values tend to be erroneous. There are two primary factors required to obtain more accurate results,

- Ejection Time
- Resistance to Blood Flow

The cuffless measurement's sensitivity including these two factors was determined by calculating the changes in the pulse and the mean arterial pressure with a 10-80% change in the ejection time for a reference resistance taken as 18.31 mmHgmin/L. From Figure 4.8 (a), it can be seen that there is a linear increase in the mean arterial pressure with increase in a person's ejection time.

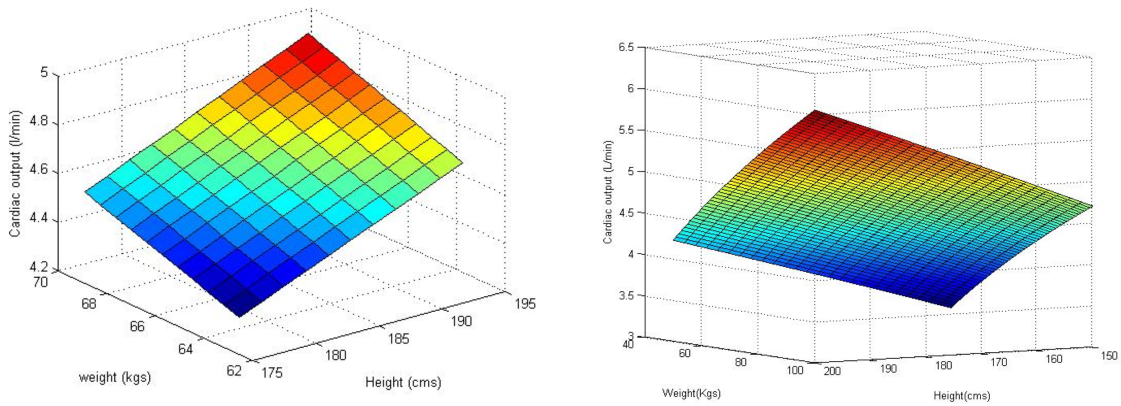


(a) (b)

Figure 4.8. *(a) Increase in Mean arterial pressure for 10-80 % increase in ET (b) Increase in Pulse pressure for 10-80 % increase in ET*

Also, Figure 4.8 (b) showed a linear increase in pulse pressure of the subject with ET. When the ejection time increased, the pulse pressure also increased according to the amount of blood in the person. A person's amount of blood is estimated by his cardiac output, weight, height, and age. Cardiac output and mean arterial pressure are directly related to each other. Therefore, an increase in the cardiac output will result in an increase of the person's blood pressure.

Figure 4.9 (a) shows increase in the cardiac output of a person with a 10% increase in height and weight. Because cardiac output increases with an increase in an individual's body-surface



(a) (b)

Figure 4.9. *(a)Change in cardiac output for 10% change in height and weight (b)Change in cardiac output for 25 % increase in height and weight of the individual*

area, the person's blood pressure also increases linearly. Also, the cardiac output increases more rapidly with height rather than with the weight of the individual, *i.e.*, tall individuals have higher cardiac output than obese individuals. Other factors which determine the cardiac output are age and heart rate. The stroke volume and heart rate are inversely proportional for same pressure conditions. When heart rate increases, blood pressure decreases. Though cardiac output is directly related to the heart rate, an increase in the heart rate will be effected by lowering the individual's stroke volume to maintain their pressure of blood flow. Age is another factor affecting the value of blood pressure, but its impact on pressure changes is low. Increase in age is effected by the increase in an individual's overall pressure only in small fractions. Also from my analysis, heart rate(HR) having a significant role in blood pressure, was mostly compensated for by a variation of blood flow resistance, thereby maintaining the individual's blood pressure in the normal range. Hence, I conclude that just having the value of HR, the blood pressure cannot be estimated.

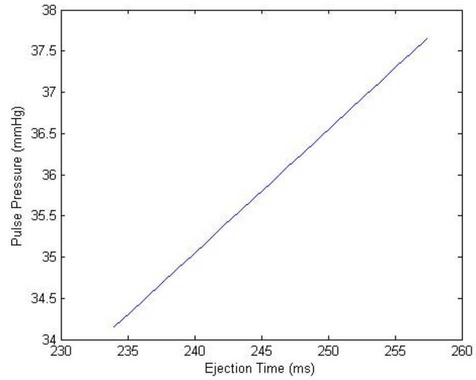
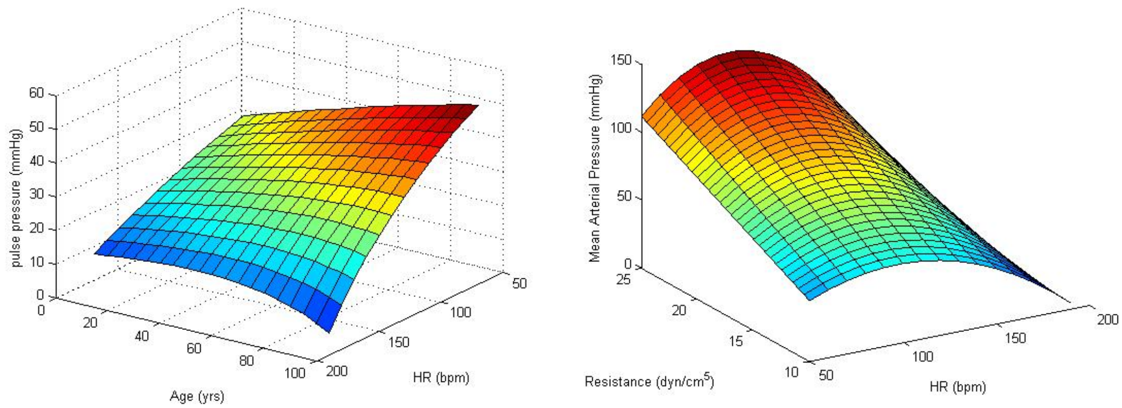


Figure 4.10. *Change in Pulse Pressure for 10 % increase in Ejection Time*

I determined pulse pressure directly from the person's ejection time based on the amount of blood flowing through the heart in each stroke. From Figure 4.10, it can be clearly seen that there is a linear increase in the amount of pulse pressure with the ejection time. The stroke volume does dependent on the ejection time, but it also relies on other factors such as the person's height, weight, heart rate and age. An increase in ejection time reflects an increase in the subject's stroke volume. I also observed that resistance to blood flow had no impact on the individual's pulse pressure. This resistance only determines the range of systolic and diastolic pressure.

The sensitivity of the recording device also must be considered while considering the sensitivity of calculations. An ideal system for cuffless estimation requires faster camera frame rates and proper placement of the finger over the lens. Obtaining an ideal result with this method will require a camera operating at 1000 fps and optimal result can be obtained with camera working at 100 frames per second. For example, the software cap on the Android platform for Nexus One [5] prevents the hardware from performing to its original capability. For instance, the hardware camera is reported to record video at 720p at a rate of 30 fps, but the recording capability has been restricted by software to 20-22 fps at 480p resolution. All the specifications mentioned here are results on testing of the Nexus one's camera under high lighting conditions. The camera's performance with respect to technical aspects will be



(a) (b)

Figure 4.11. *(a)Change in Pulse pressure for Age and Heart Rate for Normal individuals (b)Change in Mean arterial pressure for heart rate and resistance to blood pressure in normal individuals*

discussed in more details in Chapter 6. When the finger is placed over the camera lens with the led flash light illuminating the finger, rather than receiving the full light spectrum, the recorded picture received only high concentration of the red pixels, indicating a decrease in the amount of light exposed to the camera. When the picture gets darker, the frame rate decreases considerably. I have observed during the test setup that the average frame rate dropped from 23 fps to 17 -16 fps. Thus, according to my calculations, for a frame rate of 16, only 16 frames were present in a second and every frame occurred at an average time interval of 62.5 ms. Based on this estimation, I conclude that the accuracy of ejection time varied between 0-120 ms in a dataset and that the accuracy of results varied from 0 - 45 mmHg of mercury. Even the accuracy of pulse pressure ranged between 0-18 mmHg. For example, if a median error of +24 ms of ET is considered for the results, an actual 120/80 mmHg systolic pressure will show as 145/96. To eliminate this kind of error phenomenon, I took the median value of the ejection times in a dataset, thereby improving the current system's accuracy. Table 4.4 shows the proposed system's accuracy, which is better than the calculated results. Misreported values are the effect of the error conditions described above.

4.7.4. *Degree of Significance in Identification of Two Pressure Groups*

A t-test is a statistical analysis of two groups of data. t-test provides the significance of differentiation between two individuals when compared their mean values given a normal distribution. For my analysis two individuals, one with a low pressure value and the other with a normal pressure value. Additionally, I took into account, mean arterial pressure when calculating mean and standard deviation from the plot.

For analysis, ten samples were considered for each individual. The mean and the distribution of the mean arterial pressure were computed from the datasets. Figure 4.12 depicts the density plot of mean arterial pressure for two subjects.

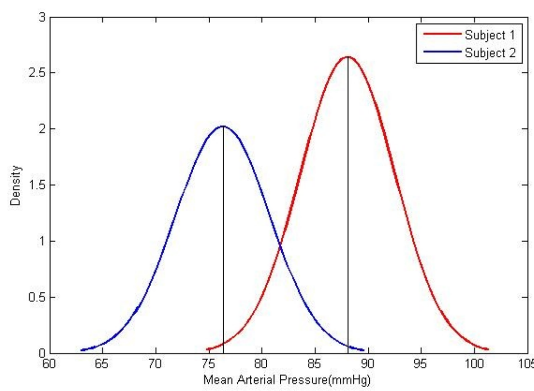


Figure 4.12. *Density plot of mean arterial pressure of two individuals. Blue and red corresponds the P_m for low and normal blood pressure individuals*

For both individuals the mean arterial pressure was estimated from the cuffless-pressure-estimation method. Table 4.5 shows estimated values of pressure.

The equations for the t-test analysis listed below were borrowed from [23]

$$t - value = \frac{Differencebetweengroups}{Variabilityofgroup} \tag{15}$$

Table 4.5. *Estimated mean arterial pressure for two individuals*

Sample #	Subject 1(Low pressure)	Subject 2(Normal Pressure)
1	80.67	91.33
2	75	88
3	80.33	87
4	69.67	91.67
5	81.33	84.67
6	77	84.33
7	69.33	90
8	76.33	80.33
9	80	87.33
10	73.33	96

$$= \frac{\overline{X_1} - \overline{X_2}}{SE(\overline{X_1} - \overline{X_2})}$$

The standard error of difference (SE) is given by

$$SE(\overline{X_1} - \overline{X_2}) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \tag{16}$$

where σ^2 is the variance for the group and n is the number of sample/members in the group. Therefore, the t-value of the groups is given by

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \tag{17}$$

The value of t should be compared with the table of significance to determine if the ratio is large enough to differentiate between the two groups. The larger the t value, the more significant the group are. However, before going to table of significance, values for the degree of freedom (df) and alpha level (α) have to be determined. The degree of freedom is given by

$$df = n_1 + n_2 - 2 \tag{18}$$

The alpha level is the probability of obtaining a t-value for individuals in the table of significance [24]. In this study, a value of 18 was obtained for the degree of freedom using the equation 18 and the alpha level was set to 0.001, which represented a probability of 99.9%. Based on the data obtained, the t-value was estimated as 5.925. The t-value in the table of significance for the selected alpha level and degree of freedom was given as 3.92, which is less than the estimated t-value. Therefore, the degree of significance was high and the differentiation between the two individuals was significant.

4.8. Cuff-less Differential Estimation

Usage of a pressure cuff for taking blood pressure restraints the users from portability. Causalities doesn't happen all the time and no-one is ever prepared for a situation with all the necessary tools. Even though the Cuff-based pressure estimation promised a significant accuracy, carrying a special device to help in measuring the blood pressure is not practical. The cuff-less blood pressure estimation discussed in the Section 4.7 did not promise high accuracy under normal circumstances. Pressure estimation needed a source of data to support and focus the values obtained from the finger pulse in cuff-less estimation, which resulted in the design of a differential blood-pressure estimation mechanism using both the finger pulse and heart sounds.

4.8.1. *Detailed Methodology*

The differential estimation of blood pressure involves the use of both heart sounds obtained from audio and pulse data obtained from the subject's finger using video. The fairly simple test setup starts with synchronization of time between two mobile phones using Bluetooth(The synchronizing procedure is discussed in detail in Section 4.7.2.). Figure 4.13 shows the overall procedure involved in determining blood pressure using this method. BPCamera records at

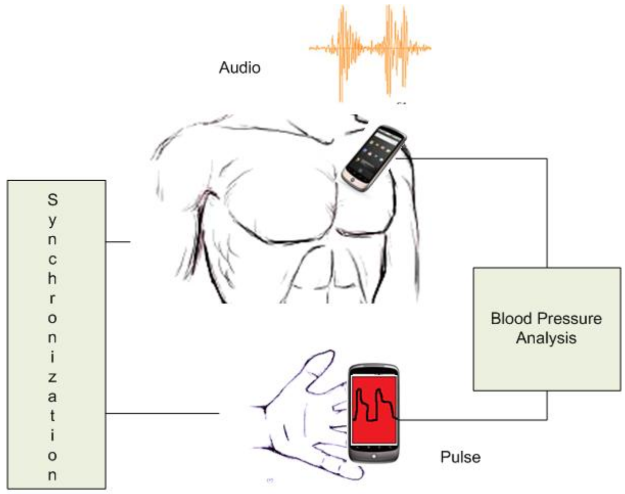


Figure 4.13. *Blood pressure estimation Test Setup: Pulse signals are obtained using the video recording app when the finger is placed over the lens and heart sounds are recorded simultaneously using the audio app by placing the other phone over the chest as shown.*

800 x 640 screen resolution with the automatic white balancing and auto-focusing disabled in the camera. The pulse in the finger was obtained from the BPCamera, the video recording application recording at 24 fps. An Led flash lights up on the application's start and stores the video recording start-time to the phone's SD card. Similarly, the audio recording application is started with the phone's microphone held close to the chest. Care has to be taken that the microphone's opening is held tightly to the skin over the chest to avoid recording external noise and unwanted audio peaks into the dataset. Blackbox V2 records the heartbeat from the chest and stores a timestamp at the onset of recording. Timestamps from both devices are used in post-measurement steps for a synchronous start of audio and video recording. Section 4.8.1.1 discusses in detail the localization of the heart beat in the chest.

4.8.1.1. *Localization of Heart Beat and Pulse.* Localizing a heart beat is a challenging task, but necessary for accurately determining blood pressure. Heart sounds are produced with the opening and closure of heart valves. Heart produces mainly 4 types of sounds [25] in one heart cycle: S1, S2, S3 and S4. The first heart sound (S1 -lub) is produced by the atrioventricular valves (*i.e.,* mitral and tricuspid), the second heart sound (S2, dub) is produced by the semilunar valves (*i.e.,* aortic and pulmonary valve). The third and fourth heart sounds are produced only in some rare conditions due to gallop. I am interested in recording only the first and second heart sound.

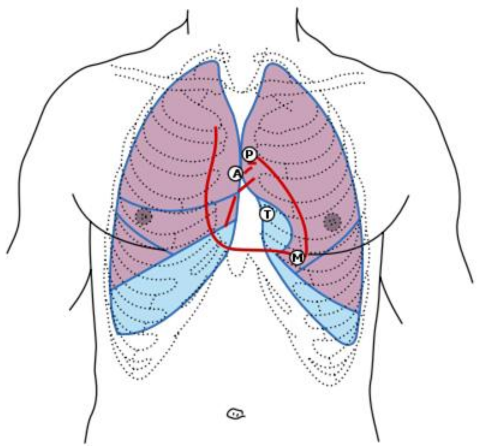


Figure 4.14. *Heart sound locations: S1(lub) is produced by mitral(M) and tricuspid(T) values; S2(dub) is produced by aortic(A) and pulmonary(P) values*

Experiments were conducted on the chest region to find the best spot for obtaining recordable heart sounds. As indicated in Figure 4.14, the four regions selected for recording purposes are where the valves are located. Based on the test results captured from the microphone, the heart sounds were heard the loudest in the pulmonary region containing the pulmonary value. Since the pulmonary value is associated with the S2 sound, in most cases the decibel level of S2 is higher than that of the S1. Table 4.6 [26] gives the location of valves in the chest with respect to the intercostal space and sternum.

Finger pulse is measured using video intensity data captured by the mobile phone's camera. Accurate localization of the finger pulse depends on proper placement of the finger over the camera lens(The proper way for holding the finger over the lens is discussed in Section 4.6.3.).

Table 4.6. *Location of valve on chest*

Pulmonary valve	second intercostal space	left upper sternal border
Aortic valve	second intercostal space	right upper sternal border
Mitral valve	fifth intercostal space	medial to left mid-clavicular line
Tricuspid valve	fourth intercostal space	lower left sternal border

As has been noted elsewhere, accurately measuring a finger pulse with the camera can be difficult. Thus, an alternate method of indicating proper measurement of the pulse in the finger is under development. However, once the heart beat and the pulse are correctly located, both the measurements have to be done in parallel to be of value. Working of the system will be discussed in the following section

4.8.2. *Inter Working of Components for Blood Pressure Estimation*

Blood pressure estimation using differential measurement starts with synchronizing both recording devices system clocks. This synchronization using Bluetooth is discussed in Section 4.8.3. Once the system clocks are synchronized, the applications for recording the data are started. The Nexus one records the video at a rate of 24 fps and the HTC Hero [27] records audio at a sampling rate of 8000 Hz. Both recording applications must start within 10 seconds of clock synchronization. Leaving the application to run for a long time, creates a drift in the system clock as the processor is responsible for maintaining the system time. When the CPU is loaded with high-priority processes, the processor cannot maintain the clock's accuracy.

The video is MP4 encoded at 800x600 resolution and the audio is a RAW PCM, working at 8 KHz following 16 bit big endian mode. In my study, all the measurements were taken for a duration of one minute with the video starting roughly 5 seconds after the audio. This step provides a crude verification of the sync and the process is optional. Figure 4.15 shows the sequential execution of events in the estimation process

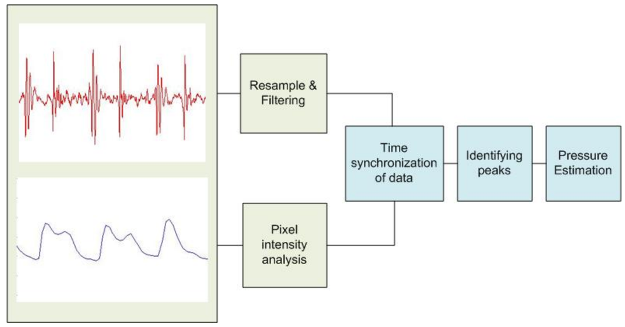


Figure 4.15. *Working of Cuff-less differential estimation of blood pressure*

The data collected were processed offline in MATLAB. The first process in the analysis is the de-noising of data and re-sampling. Audio collected from the mobile phone was passed through a 15^{th}-order low-pass butterworth filter. The frequency between 10 - 250 Hz is allowed to pass with the filter, blocking all other frequencies. Most recorded heart sounds appear within the 10-250 Hz band of frequency. For computational effectiveness, filtered audio was processed by re-sampling to a lower sampling rate of 1KHz. Similarly, the video frames were processed to give the value of red intensity. The process of estimating the intensity of a frame is described in the chapter 2.

With the datasets processed, both the audio and video data had to be synchronous in the starting time *i.e.,* adjusting the data to make them start at the same instance even though they actually start at different times. The timestamps stored on the SDCard of the two devices were used to calibrate the data's starting time. For instance, let us suppose the starting time of the audio data was S_S and the starting time of the video data was S_V. The difference in the timestamps will give the amount of time one data was ahead/lagging the other. Algorithm 2 explains the procedure followed in calibrating the two devices.

The manipulation of data was done only with the audio because, the sampling frequency for audio is 1000 per seconds, which is 1 ms for an audio frame. Thus adding and deleting data points is easier and more precise in audio than for the 24 fps Video. With both the

Algorithm 2 Calibration for Synchronous start of Data

INPUT: Timestamps of Audio and Video Start (S_S and S_V)

OUTPUT: Synchronous audio data

SyncStart = $S_S - S_V$

if SyncStart ≥ 0 **then**

Add Null data for SyncStart milliseconds towards start of the audio dataset

end if

if SyncStart < 0 **then**

Delete data for SyncStart milliseconds from the start of the audio dataset

end if

measures starting synchronously, the next step was the determination of the peaks from the data and selecting the appropriate peaks in video corresponding to the audio signal.

From the plot obtained from the audio, I could clearly distinguish the S1 and S2 sounds. However, the peak corresponding to the S1 sound in video now had to be identified. The Vascular Transit Time (VTT), the time taken by blood to travel from the heart to an extremity of the body by one stroke of the heart, was calculated by measuring the difference in time between the audio's S1 and the corresponding systolic peak in the video. Figure 4.16 shows the peaks in the audio and video data. The time of arrival of S1 and the corresponding Systolic peak are noted. For a dataset of 1 minute, I could identify a minimum of 50 - 60 data points. The median of the difference in time was taken as the VTT. VTT was identified by time difference in-between the dotted lines shown in Figure 4.16. The Systolic pressure and Vascular transit time are linearly related. The change in systolic pressure is derived from the change in VTT with respect to a reference value as described in [28]

$$\triangle SystolicPressure = -0.425 \times \triangle VTT \tag{19}$$

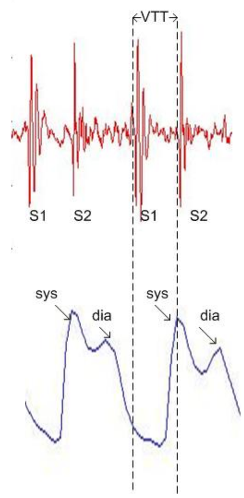


Figure 4.16. *Identification of Peaks: In the figure, sys represents the systolic peak of the pulse, dia represents the diastolic peak of the pulse. S1 and S2 are the first and second heart sound respectively. Vascular transit time(VTT) is the time difference between the origination of S1 in Audio to the appearance of corresponding sys peak in video*

Based on equation 19, the systolic pressure values corresponding the VTT can be generated. Thus, the Systolic pressure is given by

$$SystolicPressure = -0.425 \times VTT + 214 \tag{20}$$

Figure 4.17 gives the pattern of systolic pressure and VTT. An individual's diastolic pressure can be measured through calculating pulse pressure using Ejection time from the audio data, provided the value of person's age, height, and weight are known. This method was described in the cuff-less estimation of blood pressure(Section 4.7).

4.8.3. *Bluetooth based Clock Synchronization*

Clock synchronization is a common problem for computer science and engineering as it struggles with difficulty when internal clocks of several computers differ. Even when initially set accurately, real clocks differ over time due to clock drift, caused by clocks counting time

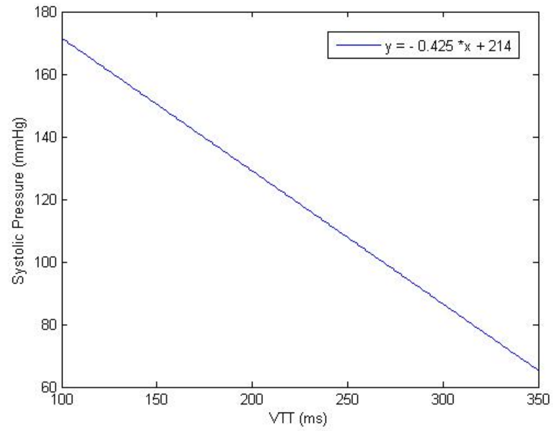


Figure 4.17. *Linear relationship between VTT and P_s obtained from the equation 20*

at slightly different rates. Even with the presence of a dedicated clock, the clock's time gets drifts. Mobile phones use the main processor for keeping track of the system clock, this increase in the processor load will cause the clock to drift more from the normal drift rate. There are different solutions to the Clock synchronization problem, the most obvious solution is the centralized system in a centralized-server environment. The problem becomes more complex in distributed environments. The most frequently used solution is the Network Time Protocol(NTP). There are various other protocols available for clock synchronization . For the purpose of Clock Synchronization, I designed a protocol which uses Bluetooth to synchronize the clocks of two mobile phones. The Bluetooth synchronization design resembles Precision Time protocol-IEEE 1588 [29] used in wired networks. IEEE 1588 was designed for local systems requiring very high accuracies beyond those obtained from NTP and GPS.

4.8.3.1. *Empirical Framework.* The phones being used in the measurement of blood pressure have to be synchronized. Bluetooth synchronization follows a Master - Slave architecture. The mobile phone which receives the synchronization messages acts as the Master and the one sending the messages acts as the slave. Both devices communicate using the Bluetooth stack, sending sync messages back and forth to adjust the clock on the system. Figure 4.18 shows the setup of the synchronization mechanism between the two phones. The phones communicate via Bluetooth using BlueZ. BlueZ [30] is the generic Bluetooth stack created for the linux distributions. In my system development, I used Android's BlueZ, a Bluetooth 2.1 compatible stack having the ability to run on any Bluetooth chipset. All the builds after the Android 2.0 is equipped with the BlueZ 4.47, which revealed Java APIs for automatic pairing and transfer of data between two devices.

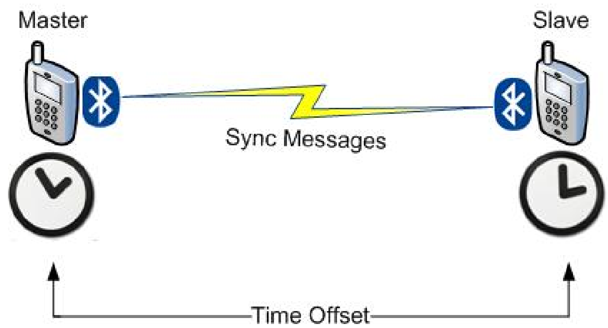


Figure 4.18. *Bluetooth Synchronization Test Setup: Master and slave device sends sync messages between them via bluetooth. Slave calculates time offset from the master clock. Slave sets the master time as its system clock*

For my experimental setup, Nexus One and HTC Hero were employed. The Nexus one, used for the video capture section, was made to act as the master device and the Hero, loaded for heart beat recording, acted as the slave. Nexus One is equipped with a Broadcom BCM4329 chipset supporting Bluetooth 2.1+EDR (Extended Data Rate technology) along with 802.11N wifi, where as the Hero supports Bluetooth 2.1+EDR along with a 802.11b/g wifi connectivity.

4.8.3.2. *Master-Slave Message Exchange Mechanism.* The over-all system follows a layered architecture for synchronization. This architecture comprises of three important layers, Application, Bluetooth Stack and the Communication layer. The application layer is the heart of the business logic for operation. (Bluetooth synchronization does not follow any pre-existing protocols and the working of the mechanism is discussed in Section 4.8.2.3.) The application layer decides which mobile phone acts as the Master and which phone to act as

Slave device during synchronization. The Device initiating the initial transfer of the sync message is chosen as Slave and the other is assigned the Master. The Master device's application layer has only the module to relay a reply message, whereas the slave device's application layer has modules to relay messages in a timely fashion and a time synchronization module to set the time on the device. The synchronization module calculates the exact time to be set on the host device relative to the master device. Bluetooth stack layer and the communication layer are the same for both the devices. Both devices have similar modes of operations and Bluetooth's performance will eventually be the same for both devices. Figure 4.19 shows the functional components of each layer.

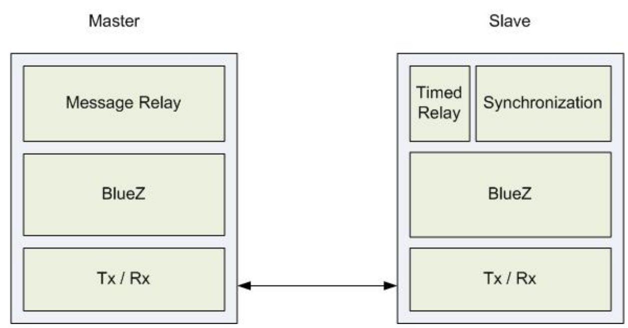


Figure 4.19. *Overall Architecture*

4.8.3.3. *Slave Device synchronization to Master Clock.* The master device's timestamp is obtained through Bluetooth and is fixed on the slave device. This simple method synchronizes time on two devices, but differential estimation of blood pressure requires time calculations precise to milliseconds. Ignoring transmission delays will lead to degradation in the system's accuracy. Hence, transmission delays have to be calculated during the procedure.

Finding the delay in transmission of messages require the time stamps of both the systems. Figure 4.20 depicts the working of the synchronization process. The first step is the standard establishment of a bluetooth connection after pairing with the device. Once the connection is established, the device initiating the synchronization process acts as the slave and sends a timestamp to the master. The timestamp sent is stored in the slave device as T_S. As soon as

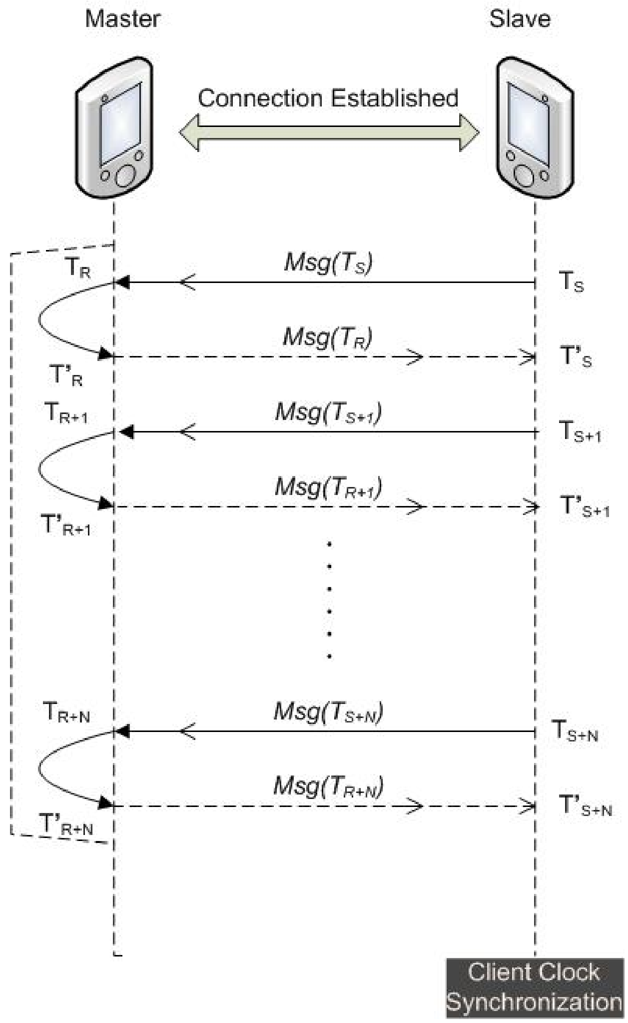


Figure 4.20. *Flow of events in Synchronization:* T_{S+i} *is the sending time in slave device.* T'_{S+i} *is receiving time in slave device.* T_{R+i} *is the receiving time in master device.* T'_{R+i} *is sending time in master device. N is the maximum number of cycles the process can continue*

the timestamp reaches the master device, it sends back its own timestamp to the slave device. This returned timestamp is recorded as T_R along with the timestamp when T_R is received in the slave device as T'_S. The process is repeated for 30 seconds and the corresponding sent and received timestamps are recorded. This process is followed by the Client clock synchronization module in the slave device. The module has four major operations: Estimating the Roundtrip Time, Estimating the Offset, Calculating the Master Time, and setting the Slave clock.

Figure 4.21 shows the components inside a client clock synchronization. The roundtrip time

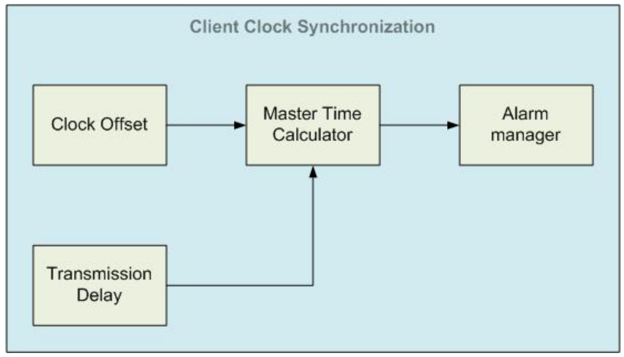


Figure 4.21. *Client Clock Synchronization*

is the difference between the message initiating time of the slave device and the receiving time of the master time in the slave device.

(21) $$RTT = T'_S - T_S$$

(22) $$T_t = \frac{RTT}{2}$$

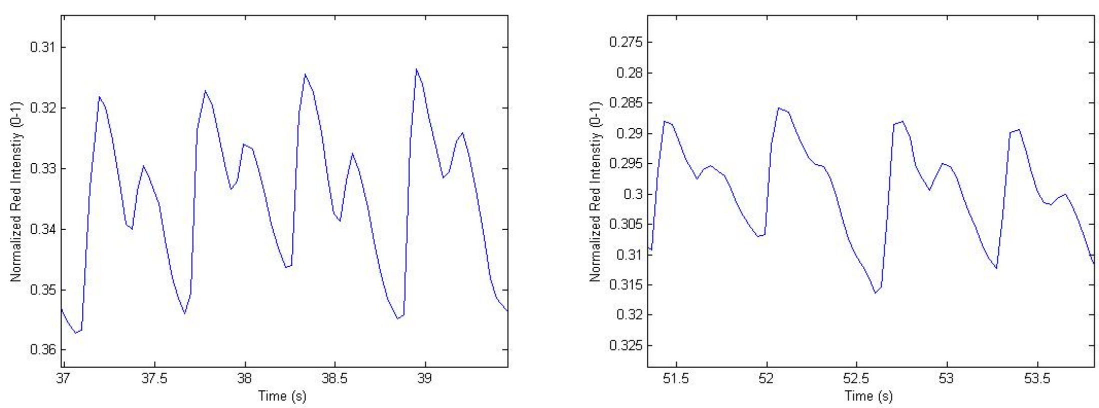
Half the value of roundtrip time gives the one-way transit time between the terminals. The offset between the two devices O_t is given by difference in timestamp of the master and slave device

(23) $$O_t = T_S - (T_R - T_t)$$

Subtracting the one-way transit time from the receiving time in the master device will produce the accurate time difference between the slave and master device. The master clock time is estimated by subtracting the offset with the slave clock. Android platform does not allow user applications to set the system time; Only rooted devices have the capability to run this application, requiring the user permission to be elevated for fixing the master's timestamp in the slave device. Thus, the Clock synchronization works between the devices using bluetooth.

4.8.4. *Accuracy with Differential Estimation*

The results from cuffless differential estimation method are promising to deliver more accurate and reliable values than the other two pressure estimation methods specified in previous sections 4.6 and 4.7. The accuracy of the results solely relies on the effective estimation of VTT using the graphs. Identification of the S1 time instance from audio was an easy task if the signal was filtered properly, but the selection of the corresponding peaks in the video intensity was difficult in some situations. The primary issue in the measurement of VTT was the availability of flat peaks for a systolic pulse in the video intensity plot. The systolic pulse in the video doesn't always have a sharp peak; sometimes it has a smooth or flat peaks leading to confusion when selecting a data point. Figure 4.22 (a) shows a graph with sharp peaks and 4.22 (b) shows a pattern of flat peaks. The flat peak pattern appears due to the sampling error in the video, discussed in the previous chapter.



(a) (b)

Figure 4.22. *(a) Sharp peak (b) Flat peak*

Table 4.7 shows a sample dataset, where the sharp peak represents the results on selection of single sharp-peaked data points and the flat peak represents the results on selection of any one data point from a series of points in a flat peak. From the comparison, it appears that selection of sharp peak gave VTT within a small range and closer to the actual pressure, whereas, with flat peaks I acquired VTT varying inside a broad range with lower accuracy than actual when compared to the selection of sharp peaks. Hence for an accurate measurement of pressure, selection of data points play an important role. Additionally, I recommend selection of data points with sharp peaks from the video data, rather than randomly selecting data point from the whole data set.

Table 4.7. *Accuracy of Sharp Vs Flat peak: The median systolic pressure was 112 for the selection of sharp peaks and 116 for flat peaks. The actual value on the blood pressure meter showed 113*

Sharp peak		Flat peak	
VTT	P_S(mmHg)	VTT	P_S(mmHg)
240	112	227	118
238	113	227	118
244	110	272	98
242	111	240	112
247	109	248	109
241	112	222	120
235	114	228	117
245	110	234	115
245	110	230	116
240	112	229	117

Table 4.8 gives the accuracy of results upon selection of sharp-peaked data. The table portrays the data analyzed from two individuals delivering the pressure measured following the proposed method along with the actual systolic pressure accured using a blood pressure

meter. The results indicate a high percentage of accuracy varying between 90 -100%. Also, my proposed method yields an accuracy above 95% in 85% of the datasets and of 90-95% in 15% of the datasets analyzed. however, as mentioned previously, the accuracy of measurement differs based on accurate selection of data points.

Table 4.8. *Accuracy of measured systolic pressure compared to the actual value on meter*

Actual Pressure	Measured P_S	Accuracy %	Actual Pressure	Measured P_S	Accuracy %
110/64	110	100	124/73	124	100
115/61	112	97.4	111/61	116	95.5
108/63	107	99.1	113/70	120	93.8
104/54	110	94.2	110/62	119	91.8
118/83	116	98.3	112/63	113	99.1
121/81	115	95.0	114/66	110	96.5
112/70	116	96.4	112/72	113	99.1
110/65	113	97.3	97/60	104	93.3

The vascular transit time(VTT) was analyzed based on the audio and video samples measured together. For every single pulse in the dataset, a VTT can be derived. In real world situations, the pressure of the person will not stay the same within a minute and varies continually throughout the day. This leads to a change in VTT value on every pulse. Even with a single VTT value I was able to determine the systolic pressure, but the probability of the results to be erroneous was high. Thus, for all the analyzed datasets, the median of the VTT was taken to give a value for the pressure. Figure 4.23 shows a plot for systolic pressure values for 40 seconds in a single dataset. The sample dataset varied inside a range of 24mmHg. The pressure value's median gave 113 mmHg which was exactly the same value from the blood pressure meter for that instance.

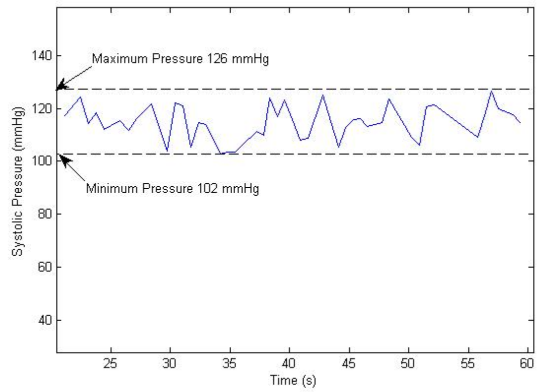
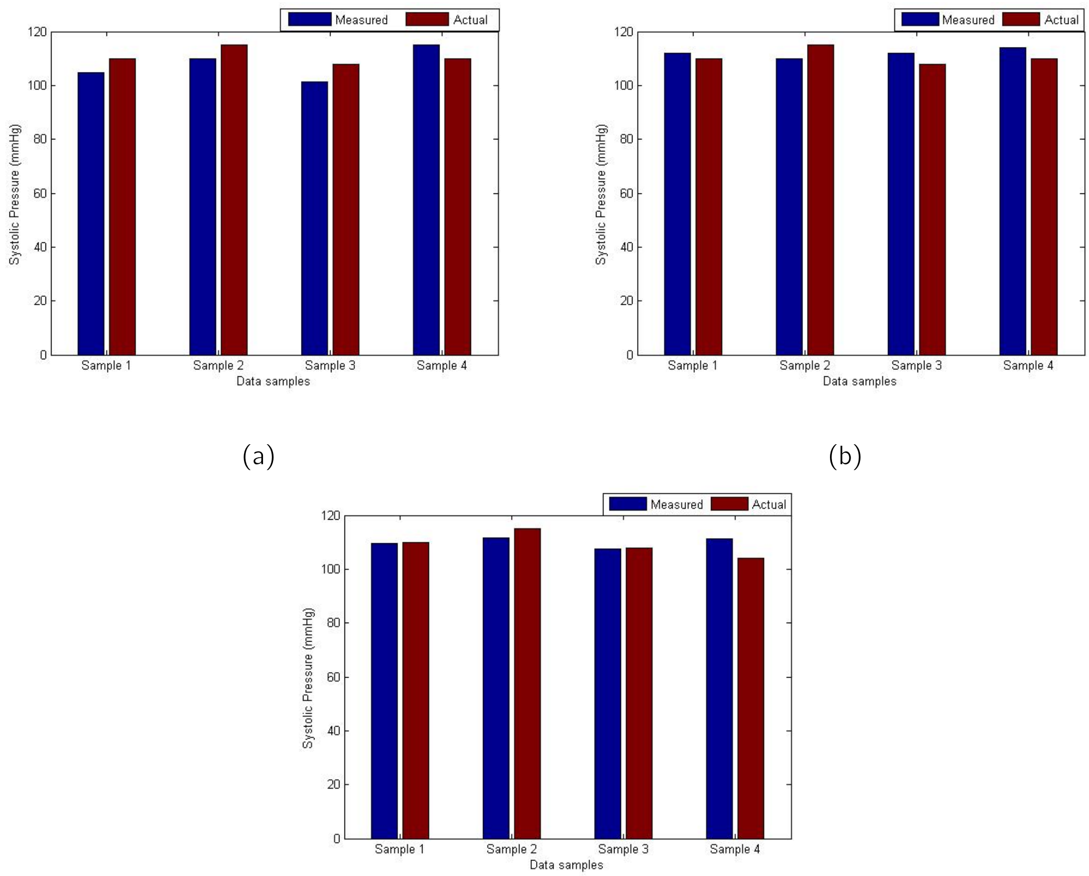


Figure 4.23. *Demonstrates variability of P_s captured using the proposed method under a single dataset for 1 minute*

The results from Figure 4.24, introduces the problem involved in the selection of number of data points. Even with the selection of a single data point, an individual's systolic pressure can be estimated, but the accuracy of the values will be at stake. Figure 4.24 contains four datasets showing the accuracy of results with the number of data points selected. Single data point in the figure represents the accuracy of estimation upon selection of a single data point from the dataset. A single point was taken at random from a dataset and compared against the actual pressure value. In a similar manner, the Five point and Ten point selection from their datasets were made to reveal the accuracy of data when compared to actual pressure values. All the plots showed results close to actual pressure, but the accuracy increased with increase in the number of data points selected.

Thus, I conclude that selection of at least 10 data points from the dataset yields better accuracy than selecting either single or five data points. Thus, the number of data points used can affect the accuracy of results. The estimation of systolic pressure from the VTT proved to be an efficient method for normal circumstances. Still, the value of systolic pressure can be determined even when the dataset has a single data. Thus, the proposed method is more robust, accurate and error tolerant than the previous methods.



(a) (b)

(c)

Figure 4.24. *Variation of Systolic pressure from actual based on Selection of number of data points from a dataset (a) Single Data point (b) Five Data points (c) Ten Data points*

4.8.5. *Impact of Age, Height and Weight on Systolic Pressure*

Sensitivity analysis of blood pressure based on Vascular Transit Time by a cuffless differential mode of estimation reveals the parameters which determines an individual's blood pressure and the accuracy of my method for determining that pressure value. In Section 4.8.2, it was noted that the systolic pressure estimate is based only on the VTT and need not depend on any other factors. In this section, I discuss on how the system's sensitivity is determined based on factors such as age, height, and weight. A theoretical model for sensitivity analysis cannot be constructed. Therefore, I match the values of datasets with the curve between systolic pressure and VTT. Figure 4.25 gives the sensitivity determined from the data for 5 subjects. The data points labeled as 'Data' refers to different individuals. Table 4.9 gives the details about the subject's age, height and weight.

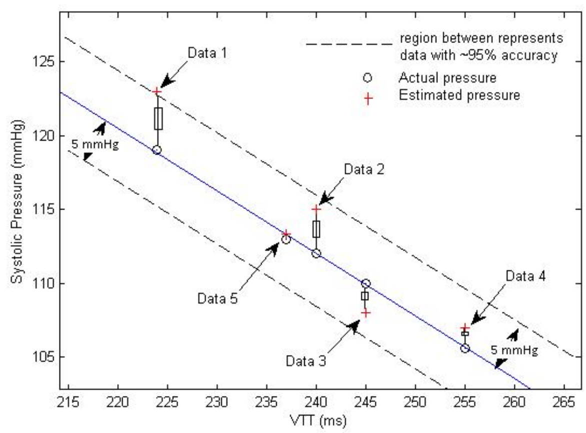


Figure 4.25. *Shows different subjects follow the same data pattern given by the equation 20. Thus the proposed method is not affected by age, height, weight of individual*

The tests were taken for subjects of varying age group, different heights, and various weight. The results calculated based on equation 20 applied to all the individuals regardless of the subject's height, weight or age, sustained an accuracy above 95% following the linear relation. Based on my analysis, age, height, and weight have no impact on the end result of proposed system. Also I observed that, for an increase of 10 ms of VTT, the systolic pressure drops linearly by 4.25 mmHg.

Based on the information from the sensitivity analysis, it can be concluded that the system will work under most conditions regardless of the physique of the individual. The estimation has not been tested with people having a history of cardiac disorders. The results and accuracy could vary for people with ailments and disabilities due to changes in their vascular system.

Table 4.9. *Subject information*

Subject	Age	Height(ft)	Weight(lbs)
Data 1	22	6.0	149
Data 2	28	6.1	142
Data 3	23	5.8	140
Data 4	25	5.8	145
Data 5	13	5.5	123

4.8.6. *Mobile hardware Performance Analysis*

The accuracy of results depends on the performance of a system under any conditions. A performance analysis could shed further light on the modular performance and blocks within the system which need improvement in the future for better estimation. At this time we feel it important to focus on synchronization using Bluetooth and hardware constraints.

4.8.6.1. *Bluetooth performance.* Synchronization using Bluetooth is the most important module in the proposed differential estimation of blood pressure. My proposed system uses two mobile phones in a master-slave relationship for collecting data. When the time difference between the mobile phone clocks increases, the probability of selecting the correct S1 sound for the corresponding Video systolic pulse decreases. If the clocks are working off sync, the resulting values from the estimation are incorrect. To estimate the performance of the bluetooth device, we assume the one-way network delay is the same for both sides, *i.e.,* the time taken for a message from a master to a slave is the same as the time taken for transit from a slave to the master. A small test for the effectiveness of the synchronization process was done. Because it was hard to have a centralized time for verification of synchronization process of two mobile phones, performance was measured in an alternative method as follows.

(i) The devices to be synchronized for the estimation were labeled as Device 1 and Device 2

(ii) Device 1 was made to act as the Master device and synchronized making note of the offset

(iii) Device 2 was made to act as the Master device and synchronized making note of the offset

(iv) When the offset value is 0, the clocks between the devices are perfectly synchronized.

(v) Steps 2 and 3 could be repeated to verify zero offset value in both directions.

The performance test was conducted on two sets of devices: Nexus one vs Nexus one and Nexus one vs HTC Hero. Hero and Nexus one does not have a common hardware configuration. Therefore, there was a tradeoff in the network delay. Table 4.10 shows the effectiveness of the clock after synchronization performed between two nexus ones. Based on analysis, I found that the lower the offset time between devices, the greater will be the effectiveness.

The offset varied ± 1 ms, which showed an effective synchronization of the clock times in both devices. The results when using a combination of Nexus one and Hero showed similar behavior, but the round trip time (RTT) between the devices is 79 ms and the offset varied between ± 6 ms. Based on my discussion in Section 4.8.5, a change in 10 ms of VTT will yield 4.5 mmHg drop in systolic pressure. Similar way, an increase or decrease in 6 ms due to offset error will introduce an error of 3 mmHg in systolic pressure. For my proposed system an offset of $\pm$ 10 ms is considered appropriate for pressure estimation. A bluetooth device has the capability to transmit data at a rate of 780 kbps according to the IEEE 802.11 standards. The effective transmission and reduction in the RTT relies on the size of messages being sent, processor capability to handle multiple requests, interference and attenuation of signal. The devices were placed as close as they can be to eliminate the delay introduced due to signal issues. Also the message size being transmitted was very small for a bandwidth of

Table 4.10. *Bluetooth Performance for Nexus One Vs Nexus One: The Device 1 acts as Master device and synchronized initially. Device column in table shows phone acting as the Slave during synchronization.*

Device	RTT(ms)	Offset(ms)
1	30	1
2	29	-1
1	30	-1
2	29	0
1	29	0
2	29	0
1	29	0
2	29	-1
1	29	1
2	29	0

780 kbps. Therefore, data loss due to retransmission, and transmission delay due to buffer overflows can not occur in the system.

Figure 4.26 gives the generalized packet structure for a bluetooth data packet.

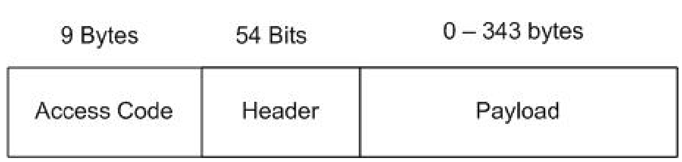


Figure 4.26. *Generalized structure of a Data packet*

During the synchronization process, a time stamp value of long data type is being sent back and forth during an instance. Considering the payload size as 128 bits, approximate size of an individual data packet was about 32 bytes. The mechanism works on simplex method sending data in one direction at a time. Hence, the message size will not create issues in transmission.

ACCESS CODE	= 72 bits
HEADER	= 54 bits
PAYLOAD	= 128 bits
Total	= 254 bits

4.8.6.2. *Microphone Performance.* The detection of heart sounds relies on the microphone's ability to pick up low-frequency, low-amplitude audio signals. Every mobile phone in the market has different hardware specifications. Two brands of phone may not have the same kind of hardware. However, a device with the best microphone sensitivity could help in recording the heart beats accurately. In this section, I discuss on comparison of microphones on three devices.

Firstly, I performed a test to determine the frequency response of the microphone on three devices, namely, Google's Nexus One, HTC Hero,and HTC Legend. The test occured in a closed environment with ambient noise and few audio activities. The three phones recorded the same audio simultaneously. The recorded audio was processed in MATLAB to derive a frequency response.

Figure 4.27 shows a flatter response curve for Hero than Legend and Nexus one. Legend has the highest power in the low frequency spectrum followed by Hero and Nexus One at the last among the three devices. Also, the occurrence of peaks in the audio data due to external disturbance was seen only in Hero and Legend, where as it was eliminated in Nexus one by the hardware noise cancelation available on the phone. Thus, heart beats can be more accurately detected using Legend or Hero than Nexus one.

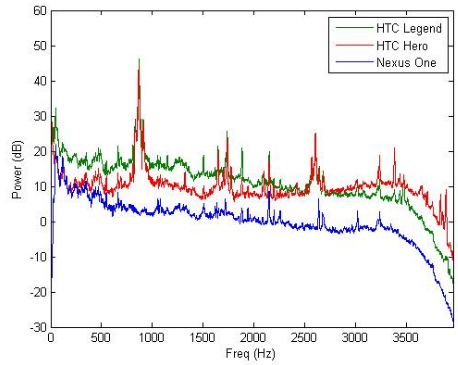


Figure 4.27. *Microphone frequency response for three different mobile phones(Nexus One, HTC Hero, HTC Legend)*

4.8.7. *Limitations of Methodology*

Some technical and social issues arise which need to be addressed when considering the use of cuffless differential estimation of blood pressure.

(i) The frame rate in video needs to be increased to achieve a more accurate reading from the proposed method. The built-in smart-phone cameras will not have a potential increase in performance any time soon. Present frame rate will be the best, any commercial user can get for the next five to six years.

(ii) The method of measuring pulses with the video intensity is native without much filtering. Hence, the appearance of sharp peaks may hardly be detected. Addition of special hardware may effectively increase the accuracy and reliability of the device, but it voids the system's purpose.

(iii) The hardware specifications for all devices are not standard. Device incompatibilities will most likely occur. For example, the LED flash available in phone may not be present in all devices as standard. Also, software version differences make the applications able to work only on specific devices.

(iv) Heart beat recording requires placement of the microphone over the bare chest. Taking measurements with the shirts on doesn't give desirable results. So, the measurement may not be possible in public.

(v) People with chronic cardiac and vascular disorder may not follow the same VTT-systolic pressure relation. Currently, because we lack of a variety of test subjects, it is impossible to arrive at a generalized relation for all categories subjects.

(vi) Calculated values of blood pressure did not match actual results, when the tests are performed with the subject lying flat to the ground.Currently, the subject needs to be sitting up right with the arm holding the camera in level to the heart.

(vii) The test setup requires two mobile phones for getting the value for blood pressure. Obviously, most users would not have two phones available during a typical scenario.

(viii) The test requires the use of rooted phones which are popular only among developers. API changes have to be made, to incorporate changes in these phone's system or the application should be developed as a core system process to be used by all public users.

4.9. Conclusion

This chapter discussed the traditional methods of blood pressure estimation. Then two methods namely cuff based and cuff-less estimation of blood pressure, both using a mobile phone were introduced. It can be said that the cuff-based estimation voided the purpose of the proposed method and resulted in an average accuracy of 70%. On the other hand, the Cuffless estimation of blood pressure only with the help of video resulted in a average accuracy of 65%. Hence another new method namely the differential estimation of blood pressure with both heart beats and video was developed and high accuracy of (90 -100 %) was obtained. Some of the technical issues of the proposed methodology and the ways of improvement have been mentioned. A performance analysis of the microphone sensitivity on multiple android based mobiles indicated that the HTC Legend was able to capture low frequency heart beats.

CHAPTER 5

REMOTE MEDIA CONTROL

5.1. Introduction

The advent of new technologies in telecommunication has led to revolutionary changes in the mobile handheld devices like PDA, smart-phones, mobile phones, tablet PCs, etc. With increasing demands in the IP technology over the traditional cellular networks/PSTNs for its cheaper cost and benefits, VoIP has been under rapid development over the past few years. Even though VoIP has so many advantages, it is still far from perfection. Despite VoIP's advantages like cheap cost, the reliability of VoIP service is still a huge question. Researchers are working hard to make the technology much better, robust and user satisfactory in every possible way. IP technology has endless possibilities for expanse; the limit is human's imagination. Any kind of service provided for a network terminal can be offered in a mobile phone with the help of IP technology. VoIP is only a part of the services offered to the mobile, considering the whole data communication. In this work, an attempt has been made to move one step towards providing a better service than the traditional cellular network by introducing remote media control during emergency situations.

Most hand held devices in the today's market have a built-in camera, speaker phone, GPS, microphone and various other sensors. During an emergency call in some disastrous situation, people panic and will not be able to convey what they intended to say. More over they are literally paralyzed to do anything. In such situations, remote media control will be extremely useful in getting precious information with the least user intervention thereby saving time and stress exerted on the user with the exhaustive questions as in the case of traditional 911 calls. A discussion of the architecture of the system, features and its overall performance will be provided in following sections.

5.2. Need for Media Control

The objective of the system is to attain control over the media elements of the phone remotely during an emergency call. When the call taker at the public safety answering point (PSAP) receives a call from a user in a distress situation, the call-taker should be able to connect with the caller's device and execute operations to control the media elements like the camera, speaker phone, microphone and other sensors.

The initial concept of creating the system is to guide the users during panic and adopting the mobile phone settings for the responder's convenience. A mobile phone is supposed to transfer audio data across a network; where as a video phone can be used to visualize the actual situation. Due to bandwidth limitations in the traditional cellular network, transmitting video might cause huge lags, delay and even dropped calls. In the IP network, bandwidth is no longer a stepping stone; now-a-days people are able to watch HD videos on their mobile IP network. Even with the availability of huge bandwidth for the network and dedicated routes, all 911 emergency situations are dealt with only audio calls. This clearly signifies the under utilization of present technology. Hence a system has been developed to integrate video and image transfer along with audio. To make the most out of the phone, Chapters 2,3 & 4 discussed a system that has been designed to provide biotelemetry and monitoring of the vital signs of a person. Figure 5.1 shows some of the media and sensor elements that can be controlled in the phone over the network.

5.2.1. *Description of Communication Components*

For the remote media control system, two clients were developed namely the mobile client (caller) and PC client (receiver). With android's open source platform and the availability of easy to use API for controls, a mobile phone powered with this operating system was used for

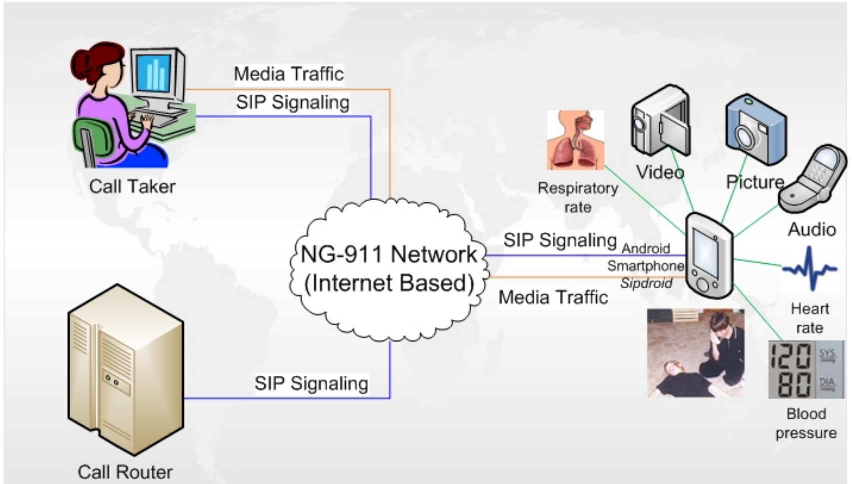


Figure 5.1. *System design of remote media control where the call taker can remotely control camera, microphone, speaker phone, built-in sensors and also obtain the victim's heart rate, respiratory rate, blood pressure with the aid of user or by-passer.*

the mobile client. The android platform is written in C and uses a Linux kernel for the core system. Android SDK provides the base for creating applications on the mobile phone using the APIs and Java. Different function libraries were used and custom java packages were imported to develop applications. The mobile client runs on most of the Android powered devices and tested on T-mobile G1, HTC Hero, HTC Legend and Nexus one. Google code has hosted an open source project; 'Sipdroid' [31] for creating an SIP/VoIP application for the Android platform. Android platform does not have API's to support VoIP applications, hence a team of developers have created a VoIP application for the mobile. This open source application was used as a base to build on the video, image transfer and media control services.

The PC client was developed over pure Java and it extends the support from an open source SIP stack called 'MjSIP' [32]. Both the mobile and PC client uses the same SIP/RTP stack for communication. MjSIP is a Java SIP/RTP stack which runs on any operating system. It makes use of the Java direct audio and Video Conferencing tool (VIC) [33] for receiving video from the mobile client. The requirement of the system was so specific in the direction of communication, hence most connections will be unidirectional for bandwidth conservation. Figures 5.2 and 5.3 show the user interface of the Mobile Client and the PC client.

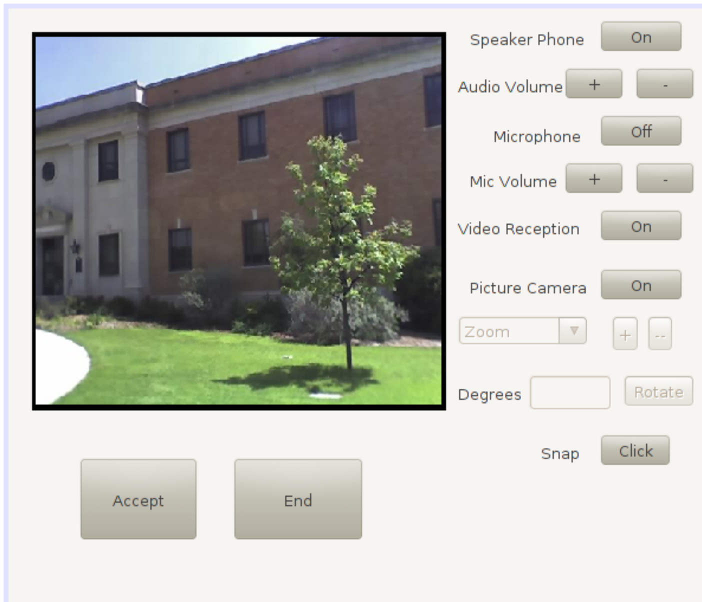


Figure 5.2. *PC Client UI*

The mobile client can make a SIP call, in full duplex voice transfer, but has a unidirectional video transfer; it cannot receive a video stream. Some of the capabilities of the mobile client are to call transfer, support for both tcp/udp protocols for signaling etc. Meanwhile, the PC client can only receive calls and access to make calls is restricted. It can transmit and receive voice at the same time, but it can only receive video. Another important feature of the client is to capture images instead of video, as videos are more bandwidth consuming. In most cases a video stream is not required by the receiver, since an image transmission will provide more information than a low quality video stream. Moreover the user will have the option to decide their level of privacy during the call.

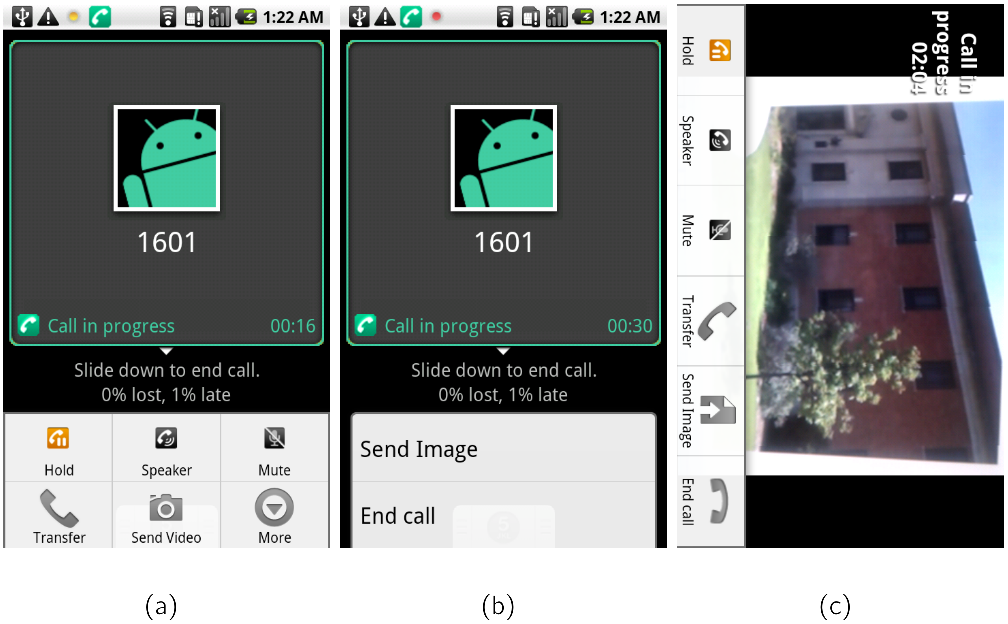


Figure 5.3. *(a) Call Options on phone (b) Image transfer option (c) video transmission*

Table 5.1 compares the technical features in the mobile and PC client. The mobile client runs on HTC Hero with Google, running Android OS 1.5. It runs on a 528 MHz Qualcomm MSM7201A processor with 512 MB ROM. The Hero has built-in GPS, WiFi and a 5-MP camera. The mobile client runs on Wi-Fi at wireless-G speed and connects to the 'nslsip' server. The video is transmitted as dynamic payload from the mobile client to the PC. The application is also tested over T-Mobile G1 and Google's Nexus one. G1 doesn't support many camera features like zoom, brightness adjustment and much more. It is recommended to use HTC hero or better. The data communication is carried over directly between terminals and does not use proxy server. The PC client runs on both Linux and Windows operating system with JVM 1.5 or higher. The current setup uses a Core 2 duo 2.53 GHz machine running Ubuntu 9.10 with open-jdk 1.6. It uses java sound for audio; it also has support for JMF /citejmf and RAT for audio transmission and reception. Video reception is taken care by VIC in the PC client, but it does not support the reception of dynamic payload types, hence a JMF was used which partially supports H263+ with minor tweaks to the codec information.

Table 5.1. *Feature comparison of the mobile and PC client*

Media feature	Mobile Client	PC Client
Audio	Full Duplex	Full Duplex
Audio Codec	G.711	G.711
Audio Encoding	PCM A-law 8Khz Mono	PCM U-law 8Khz Mono
Audio bit rate	64 kbps	64 kbps
Video	Send Only	Receive Only
Video Codec	H.263+	N/A
Video Encoding	3GPP. 90Khz (Android OS)	N/A
Video bit rate	Variable	N/A
Image	Send Only	Receive Only
Image encoding	JPEG	N/A

5.2.2. *Description of Mobile Client*

SipUA is the application developed for the IP based telephony service for android mobile phone. The user interface consists of a dialer and a window for changing the setting like the username, password, server Address and the port to which it needs to connect. For now the application can access the internet over Wi-Fi, 3G and Edge. For experimentation, Wifi connection was used for accessing the server. When the information is entered by the user, the User Agent (SipUA) registers to the registrar server, which will be notified by a small light in the Notification area of the Android mobile phone. When a number is dialed, the telephony manager is called to pop up the default dialer interface for calling. SIP messages are passed through the MjSIP SIP/RTP stack which provides the methods for VoIP calls. The audio from the microphone and the receiving packets are ported by the Sipdroid to the ear piece through the media management. When the Back key is pressed, a disconnect message is sent and the call information is stored in the phone's call Log. The information for the call log is formatted by the SipUA, by taking the user name of the number called/received as the display name and the entire SIP URL as the user number, along with the duration and direction of calling. On event of incoming calls, the calls are listened by SIPUA which starts the ringer Manager. When the user presses the connect button, the call will be accepted and the RTP traffic starts to flow through.

5.2.3. *Message Exchange between Mobile and PC Client*

Figure 5.4 shows the basic architecture of the remote media control, where the caller and callee establish a session using SIP to send control messages over the network to access the device. The default SIP stack provided by sipdroid, did not provide the extension for sending message, hence a SIP extension for instant messaging (RFC3428) [34] was implemented into the system. A new MIME type for the control messaging was created to the system called as 'application/control'. The creation of new MIME type avoids the interference of plain text messages with the control messages.

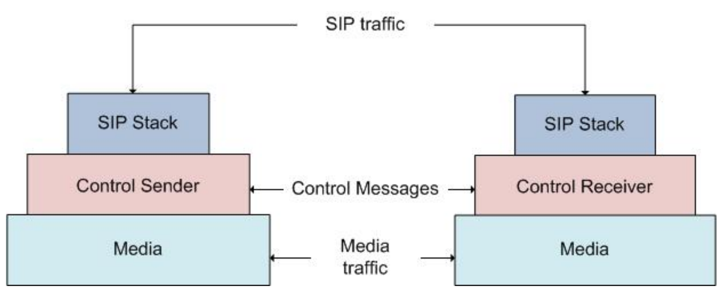


Figure 5.4. *Message exchange at each layer between Mobile and PC Client*

The control messages are sent from the PC client to perform an action in the mobile client. The control sender module interacts with the UI and forms the control code for transmission. As soon as the control messages reach the mobile client, they are parsed and sent to the appropriate media sections for execution. An acknowledgement is sent back to the PC client on successful delivery of the message. Figure 5.5 gives the layered architect of Mobile Client.

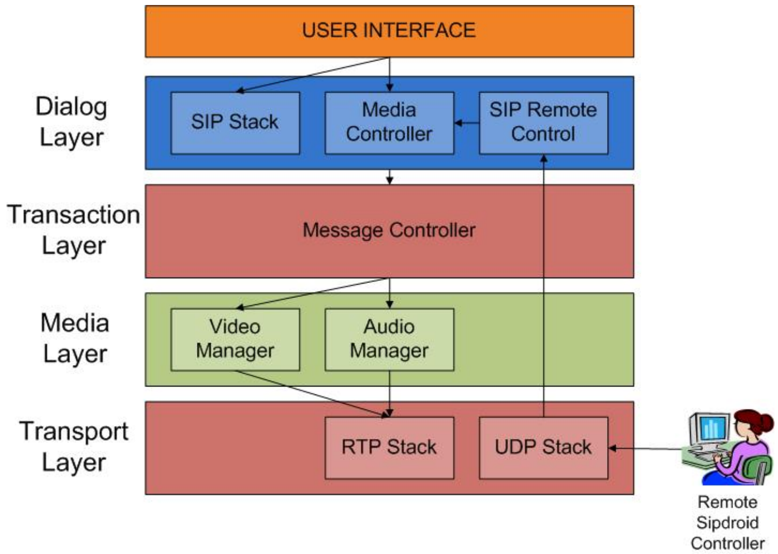


Figure 5.5. *Architecture of Mobile Client, where the remote control module is directly linked with the media controller to perform I/O operations with hardware*

The architecture gives a hint about the components of the system which are working together to complete the product. The user interface forms the foreground of the application and it provides with the options and interface to make calls for the client. The user interface lies in the application layer of the stack, which gives controls to the user for operation. The Android platform doesn't come with a SIP stack of its own and neither does have a RTP stack for transport. Hence the MjSIP stack has been used for the session and transport layer. MjSIP is a layered architecture and complies with the RFC 3261 [35]. It basically has three core layers: Transport, Transaction and Dialog. For the purpose of media control, a new media layer has been introduced to access the API level media components of the phone through its SDK.

The initialization of SIP stack is accompanied by the system configuration, initialization of logging activities and other resources like, the sip providers, listening agents. The dialog layer helps in managing the dialog objects and maps the transaction to the corresponding dialog objects. The transaction layer creates and manages transactions by mapping the messages to their transaction objects. Each transaction will maintain its transaction states, send and receive requests through the low lying layers.

The media control layer is introduced in the dialog layer, where control messages will be received and parsed to give the control codes. The commands to be executed directly to the hardware during a communication are dispatched by the media control layer. All the sensors objects are derived to this layer, thereby making it easier for directing commands to the hardware like the audio and video.

The MjSIP stack cannot provide any default controllers for the audio and the video media; hence a separate layer to access the devices using the OS API's was created. The packet format information was added from the transaction layer to the hardware data. The audio engine contains the audio encoder and uses a PCM audio at the rate of 8Khz. The layer is responsible for all the audio and video operations and contains the modules for the speakerphone and the microphone. Video transmission is done by using the media recorder without the audio data. The RTP stack of MjSIP provides the APIs to create audio RTP packets, where as it doesn't have any API for video. Hence the encoding was done by hardware and the RTP packetizing for both audio and video was software controlled at the transport layer.

5.2.4. *Control Message Transfer Process*

The initialization of the SIP session is the same as the normal procedure, where an invite message is sent by the caller and both the parties will negotiate upon the media codecs to be used for the session. Once the negotiation is complete on accepting the call, the session is established and the data transfer starts between both the ends. In the media control system, the remote media controller components will be enabled on call connection, provided the receiver can support it. When there is a need for remote control of the phone, the call taker can send the control messages specific for different operations. A sample of the control message format is given

```
MESSAGE im:user2@domain.com SIP/2.0
Via:  SIP/2.0/UDP user1pc.domain.com
From:  im:user1@domain.com
To:  im:user2@domain.com
Call-ID: asd88asd77a@1.2.3.4
CSeq:  1 <MESSAGE>
Content-Type:  application/control
Content-Length:  18
```

Once the control message is received at the mobile phone, the message is parsed to recover the control code and specific operation is performed. On successful completion of the operation, an acknowledgment is sent back to the call taker. If the message delivery fails, it will be retransmitted and follow the standard flow control mechanism.

5.3. Image Transfer

Bandwidth requirements placed a rigid blockhead for video over IP for a long time. With the availability of high speed data services, video transmission is no more a huge problem to discuss, but there is still hesitation from users in using the service because of its primary disadvantages like lack of privacy, bandwidth overhead in case of multiple video communications. Hence a system to transmit images rather than video transmission during SIP communication has been developed. Image transmission has its own advantages like resource conservation, user control over the scenes transmitted. To determine the optimal factors for the image transmission a series of experiments were conducted to identify some of the factors like resolution, zoom level, brightness, etc.. In addition, a description of architecture for image

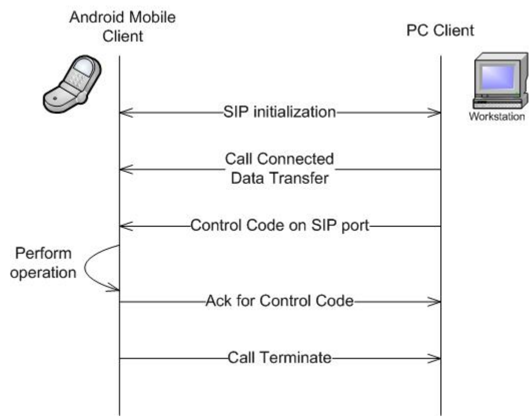


Figure 5.6. *Flow of events between the two clients: Firstly the session is initialized during the SIP initialization followed by data transfer when the session is established. Upon session establishment, the media control module will be active to send control messages from the PC client. On receipt and successful execution of the control message, an acknowledgement is sent back from mobile client*

control from a remote end is provided. With the study of these parameters, the conditions for optimal transmission of images to the call taker in an emergency calling situation can be determined. In the next section, an added functionality that was introduced for sending images from the camera to the call-taker during the call will be discussed.

5.3.1. *Description*

As described earlier in section 3.2.1, the phones run on java at both ends and incorporate modifications as per the needs. One of the enhancements for the client is the ability to send pictures to the other end for conserving bandwidth and save resources. The mechanism for sending the images/files during a SIP call is described in the RFC4976 [36], which sends the file by using a MSRP(Message Session Relay Protocol) relay server to initiate sessions for file transfer and help in the NAT traversal for file communication. Since a reliable communication is required between the end clients without much hassle, FTP file transfer was opted for, where the data is pushed by the client to a ftp server and read at the other end. This mechanism is so simple to implement and at the same time, much more robust to have a guaranteed file transfer.

There are some restrictions in the directional usage of image transfer. The mobile client can only send the images after capturing and cannot receive any images. The PC client can receive the images and send an acknowledgement for receiving the image. Hence both devices work only in half duplex mode. There are many options/parameters which can be controlled for the camera in the image mode, rather like the video recorder discussed earlier. The image mode is very similar to that of a digital camera in most of its functionalities. All the camera parameters can be controlled over remote and some functionality are shared by the end user for privacy protection.

5.3.2. *Camera Specification*

Since both the HTC hero and the Nexus one contain the same camera controller, either of them could be used. The cameras support a maximum resolution of 4.92 Mega pixel and has the capability to autofocus on objects while capturing images. It has an optical zoom level of 1x and a digital zoom level of 4x. JPEG is the only image encoding available on the phone, but YUV / RGB uncompressed color data can be obtained by the help of capturing with Native development.

5.3.3. *FTP Based Image Transfer using SIP*

Figure 5.7 shows the basic architecture of the system and the components involved in Image transfer. The file is transferred to the other end with the help of file transfer protocol in the port 21 and the SIP signaling/initialization takes place through the registrar/ SIP proxy. Once the SIP connections are established the clients can communicate with each other directly. The control messages are transferred from the PC Client to the mobile client through the instant messaging protocol. Also, there are many remote control options added for the camera module to be controlled.

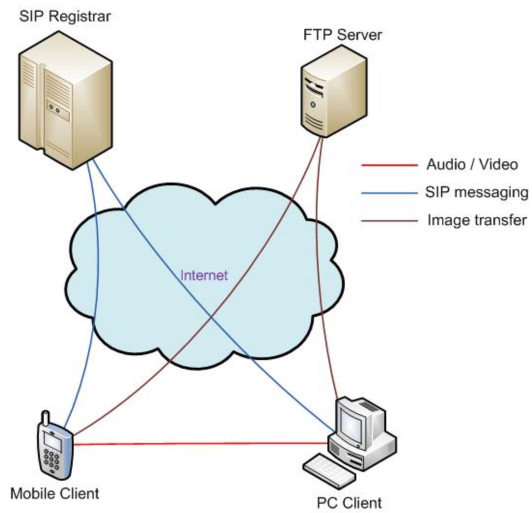


Figure 5.7. *Architecture of Image transfer Mechanism: FTP server stores the images sent by mobile client and PC client reads them with the file name sent through the sip messages*

Figure 5.8 represents a small module stripped from the sipdroid application for sending the images to the server from the Mobile phone. It basically contains three major components, File namer, FTP Sender and the camera module. Whenever the user wants to send the image to the call taker, they can start sending by clicking on the option in the menu of the application. Else the call taker can remotely turn the camera ON and send them. The file namer is a module which generates the name for the image file to be created. This module is important as it generates a random filename and conveys the file name to the destination/receiving side for reception. This provides a mechanism of authentication since, the file name will be visible only to the people involved in the communication. The FTP sender module creates a client side socket to start a file transfer to the server. The FTP sender uses standard authentication mechanism to open the FTP socket. The camera module includes a picture taking operation along with features for controlling the image factors like auto focus, brightness, contrast, sharpness adjustments, etc. Most of the camera features are embedded in the camera module.

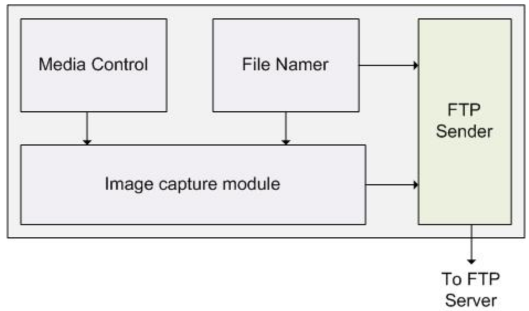


Figure 5.8. *Image transmission Module: Media control enables the image capture module. When the image is taken, the File namer generates a random sequence of name for the file and uploads to the ftp server.*

5.3.4. *Event Flow for SIP Assisted FTP Transfer*

The working of the image module is simple in the sense; it uses the idea of MSRP relay servers without handling all the negotiation to accept/reject the transfer. The image transfer mode starts only with the establishment of the voice connection between the two SIP clients. Once the call is established the modules are enabled and when the module is started, it displays the preview on the user screen. The images can be taken either by the user or taken over the remote by the click of a button.

When an image is taken, the filename is generated for the image and saved in the mobile phone. The FTP uploader establishes a connection with FTP server after the authentication mechanism. Once the authentication is done, the image is transmitted to the FTP server and gets deleted at the mobile phone. The file name is generated randomly to avoiding trespassers to guess the files from the server, in case they are able to break in. The random file names provide a certain level of privacy from attackers. As soon as the mobile client receives the notification for file completed, the name of the file is transmitted across the SIP clients using the Instant messaging protocol. The PC client having the filename can directly request the FTP server and download the image file after establishing an authenticated session. Thus

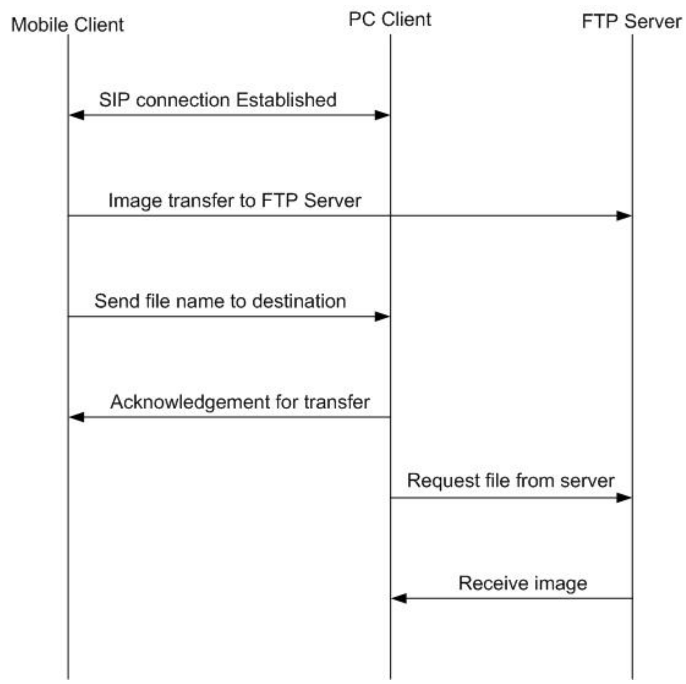


Figure 5.9. *Flow of event for image transfer*

the image is completely transmitted to the other end without any problem. The Figure 5.9 gives the sequence of events for a complete image transfer. There are some of the features that could be controlled remotely.

Table 5.2. *Range of values that can be set by from remote by media control*

Feature	Range
Zoom	1-40
Brightness	0-10
Contrast	0-10
Sharpness	0-30
Saturation	0-10
Rotation	0-360

The Table 5.2 shows the features that can be controlled from remote and also the mobile client can allow the caller to zoom in/out and take picture from the camera for transmitting. Whenever the camera is taking a picture, it will auto focus itself as the focus of the camera can't be controlled manually.

5.3.5. *Impact of Media Features on Image Quality*

The size of the image is directly related with the quality of an image. For an image of the same resolution and fixed scene, the quality of the image will increase with increase in file size of image [37]. The quality of the image increases with the size, because increase in quality generally means increase in amount of spectral information present in a pixel. When the number of pixels per square inch increases, the data volume increases resulting in increase in the size of image. In this section, it is assumed that the file size is directly considered with the image quality and demonstrate the effects of various parameter of the camera over the quality of the image.

5.3.5.1. *Perceived Quality to Image Resolution.* In this experiment, the size occupied by the file with changing resolution was measured. The test was conducted by keeping the mobile phone attached to a fixed stand and varying the resolution given in number of mega pixel of the camera, the image size varied from 0.3 mega pixel to a maximum of 5 mega pixel (actually 4.92 mega pixel).From Figure 5.10, it can be observed that file size and the resolution are linearly related. The increased image resolution, results in the improvement of the image quality with more number of pixels per square inch, thereby leading to high spectral intensity of the image.

A continuation of the first experiment was to identify the time taken for the image to be uploaded in the ftp server. This mainly depends on the available bandwidth and the link speed available to the system. For our experiment, a 54 Mpbs link speed was used for the mobile phone over Wi-Fi, which gave a near zero latency for the transmission. From Figure 5.11, it

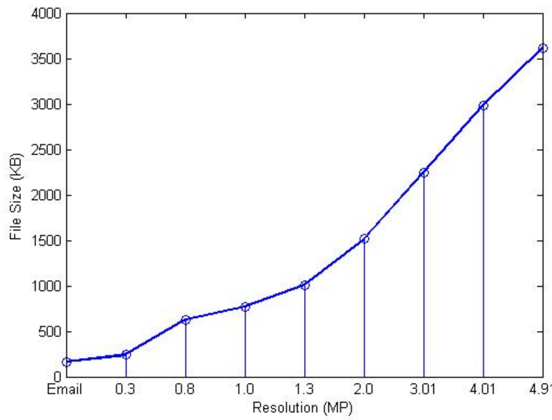


Figure 5.10. *File Size Vs Image resolution*

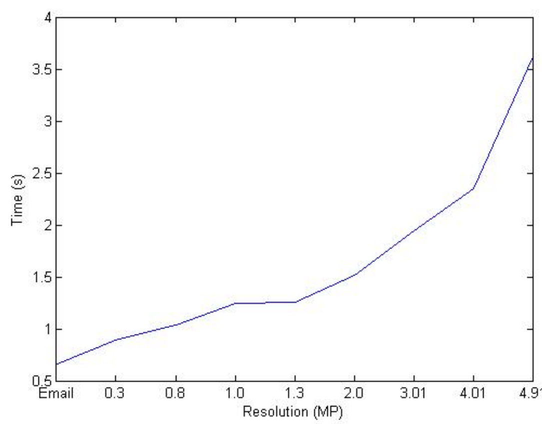


Figure 5.11. *Transmission Time Vs Image Resolution*

can be observed that the transmission time increases with increase in the resolution of the image. It means that, the file size increases when the time taken for the system to transmit increases. It can also be observed from the graph that there is a gradual increase in the transmission time for a certain size, followed by rapid increase in the time. This is because of the buffer size used in the transmission.

5.3.5.2. *Perceived Quality to Brightness, Contrast and Sharpness.* Some experiments were conducted to identify the factors that could possibly affect the quality of the image upon control. In all the experiments, the mobile phone was fixed and stationary, under the same environment. There were no changes to the scene except, the variable factor that was considered to test. All images were taken at the email size of 512x384 for the best performance, in terms of image transmission. Now the increase in zoom of the camera with image quality is considered. From Figure 5.12, it can be seen that the file size of the image decreases. Hence the quality of the image decreases on zooming.

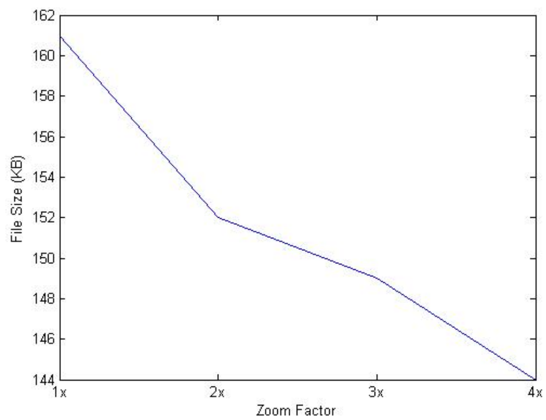


Figure 5.12. *File Size Vs Zoom factor*

The brightness parameter of the camera was varied, keeping all the other factors constant. From Figure 5.13, it can be observed that the image quality is degraded in low intensities and increases gradually with increase in the image brightness. It can also be seen that, the quality remains constant after a certain level of brightness. This proves that the camera works better in medium and bright lighting conditions than the low ones.

Contrast of an image is the difference in visual properties that makes an object distinguishable from other objects and the background. It brings about many changes to the perception of the image by an individual. For observing the quality of the image, its contrast was varied between -3 and 4. From Figure 5.14 it can be concluded that the image quality was degraded when the contrast was high or low. However there existed only one optimal contrast value for maximum quality.

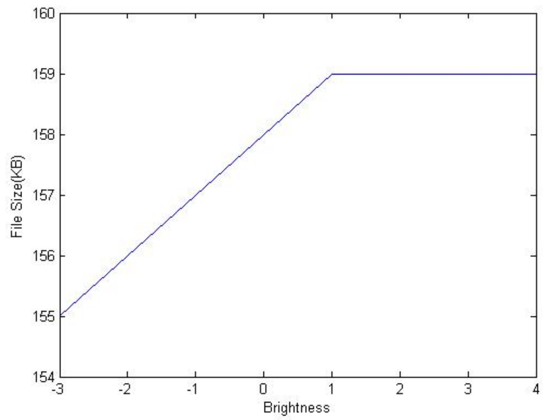


Figure 5.13. *File Size Vs Brightness*

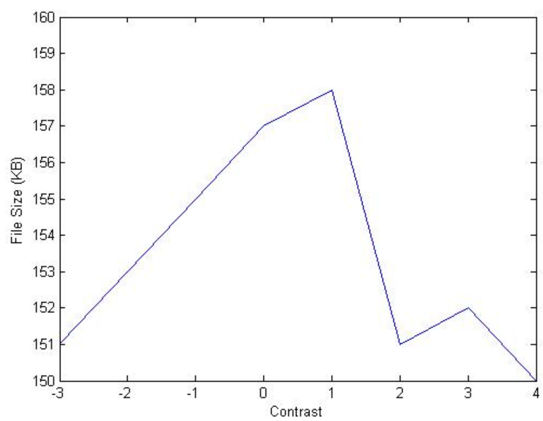


Figure 5.14. *File Size Vs Contrast*

Sharpness of an image contributes to the demarcation of an object from that of the other objects in the image. Blurriness increases when the image is not sharp and tends to overlap pixels on the boundaries of other objects. Based on the observation from Figure 5.15, it can be seen that the quality of the image increased as the sharpness increased and remained constant above a certain line. This is because, the sharpness cannot impact an image after a certain limit and the increase in the sharpness beyond a threshold will result in the increase of every pixel detail of the image, adding up noise as a whole. For the use of our mobile camera , the sharpness cannot be increased beyond a certain desirable limit.

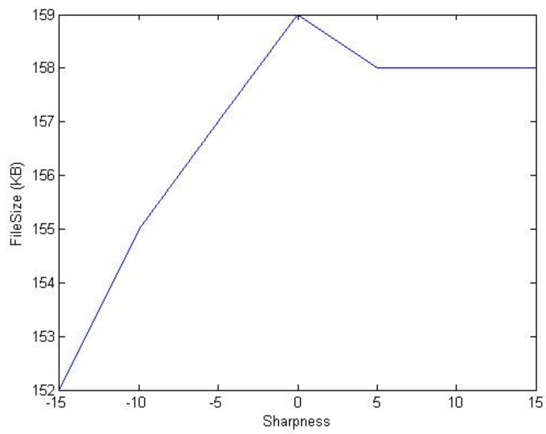


Figure 5.15. *File Size Vs Sharpness*

5.3.5.3. *Perceived Quality to Camera Exposure.* The quality of the picture depends on the type of light source and the amount of light exposure it receives. A simple experiment was conducted to identify the best light source for the mobile phone (HTC Hero). Based

Table 5.3. *File size under different Lighting Conditions*

Light Condition	File Size(KB)
Dark (No light Source)	29.2
Indoor Natural Light	67
Indoor Low Light	72.4
Indoor Bright light	118
Indoor on Light source	132
Outdoor low light	169
Outdoor medium light	197
Outdoor bright light	202

on the results tabulated in Table 5.3, it can be seen that the file size/ quality is more in outdoor lighting conditions than indoors. Also, there is a better performance of the camera performed in the presence of a brighter source than gloomy ones. The tests were conducted in darkness, indoors with a single incandescent light source and a natural light source from the windows and outdoor conditions. Hence it can be concluded that the optimal quality can be obtained in outdoor environments under adequate lighting conditions.

5.4. System Performance

Remote media control is an useful tool for controlling the media features of the mobile phone in case of an emergency call, presenting less user intervention. Emergency calls are time sensitive and require reliability of information or data. If the receiver does not hear anything at the other end, then the caller will be in a dire situation. In the proposed system both audio and video were used at the same time leading to some rigid bandwidth constraints. Since reliability and clarity of information carried to the other end are given much importance, some experiments were conducted to analyze the bandwidth usage of the system and the utilization of the phone CPU at various instances.

5.4.1. *Architecture of Utilization Monitor*

For measuring the performance of the system, two separate routines were developed to monitor the activities in the mobile phone, like monitoring the network and CPU usage. The routines run as service in the background, continuously recording data. Even though android tool kit provides a benchmark tool to measure the overall performance of the system, we created system services to monitor every individual process in the phone for better clarity of the results

Figure 5.16 shows the basic architecture of the utilization module. Android OS is very similar to Linux operating system and every process running in the system is actually a file. Android operating system collects statistics for the CPU utilization for individual processes and stores them in separate files under /proc/stat. Similarly it also records the network utilization, total bytes in/out, from which the network data rate was determined. The utilization monitor

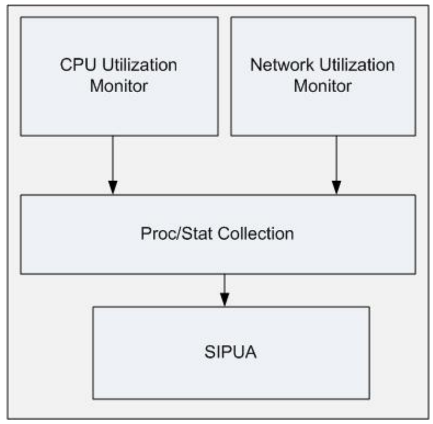


Figure 5.16. *Architecture of utilization Monitor: Android system collects statistics about processes. Utilization monitor parses the collected information to retrieve data for desired process.*

services run in the background, constantly reading the file and computing required result from the statistics file. Since the module runs as service, constantly probing the file, it requires more CPU than most other processes. In Android platform, when a process with higher priority enters the processor all the other services stall and continue after the processor is less occupied.

The test setup comprises of the mobile and the PC client registered under the same registrar server. The mobile client is running the utilization monitoring services in background. The PC client was replaced with IP hard phones that support dynamic payload data to test the performance under different hardware equipments. Figure 5.17 shows the test bed setup while measuring the performance of the system.

5.4.2. *Network Utilization of SIPUA*

Once the application is installed, data from the files is recorded by the monitoring service and stored in the user given location on the SD card. The CPU utilization monitor is designed to collect only the CPU usage of the sipdroid process and all its child processes. As soon as the mobile client registers to the server creating a process, its utilization values will be

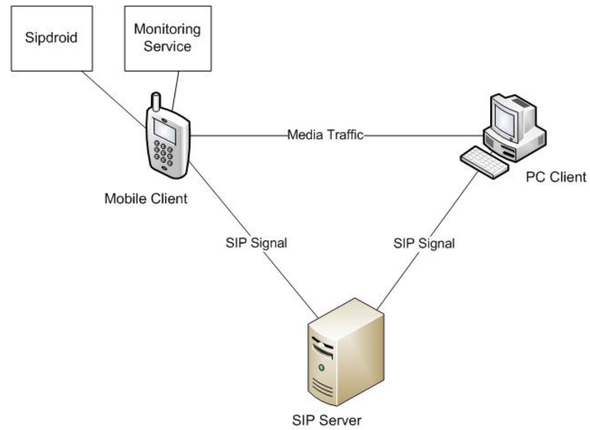


Figure 5.17. *Test setup consists the PC and mobile SIP client registered through the same SIP server under a local wireless network. The mobile phone runs the monitoring service in parallel to the client application.*

recorded to the file. In this experiment the total network rate at any given instance of time is computed using equation 24,

$$Bitrate = \frac{BitsOut_t - BitsOut_{t-sampletime}}{sampletime} \quad (24)$$

Where t is any time instant and sample time is the interval between two data recordings.

Figure 5.18 shows the bit rate when the application is idle *i.e.*, no media transfer across terminals. The small amount of bits sent is due to the keep alive mechanism in the client to probe the server at constant intervals of time. It can be observed that, there is almost negligible amount of data transfer when the client is idle and the small amount of bit rate is due to the transfer of around 225 bytes of register message.

The experiment was extended to find the bit rate of the audio. Since a static payload PCM audio was used, data transfer occurred even if no voice signal was generated. In this experiment, on the onset of the 30th second, the speaker was turned on to detect any changes in bit rate. The results showed that there was no influence over the bit rate with the change in the microphone gain.

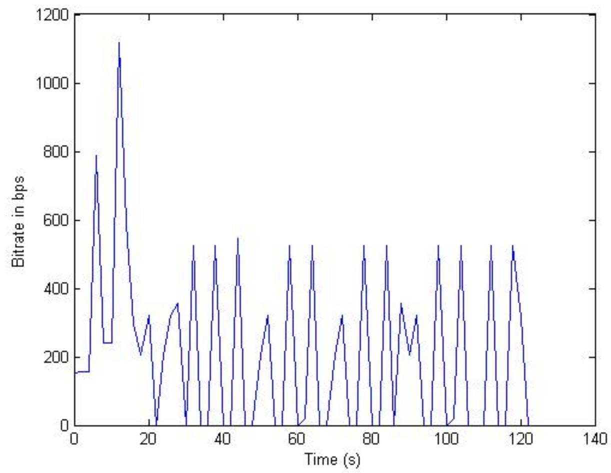


Figure 5.18. *Network utilization with Client is idle*

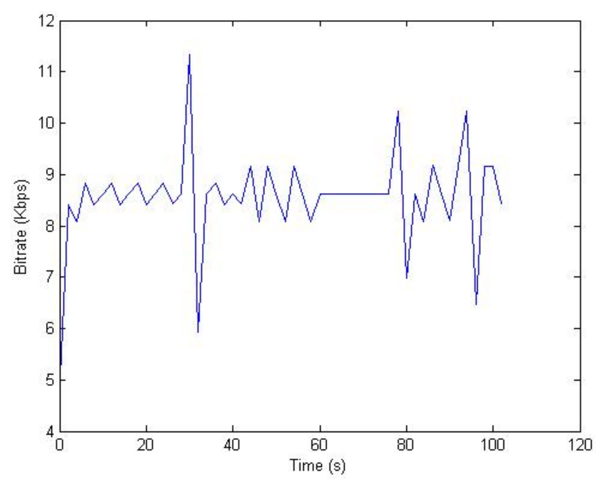


Figure 5.19. *Network utilization over a Voice call from client*

From Figure 5.19, it can be concluded that there will not be a change in the bit rate unless the codec is changed in the system.

The experiment was further extended by transmitting a video with the camera moving along with its bearer. The application had its speaker phone turned on for audio along with the video.

As the codec used was variable bit rate; the packet size varied from 240 bits to 1458 bits depending on the current frame size. Due to the constant motion in video, it can be observed

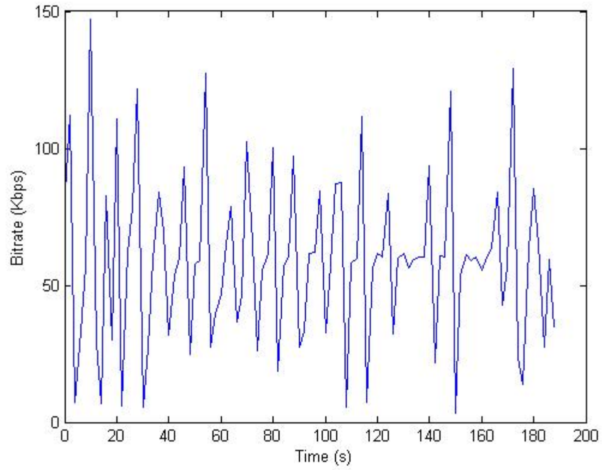


Figure 5.20. *Network utilization over a Video call under constant motion of the scene shows high frequency in bitrate changes over time*

that the frequency of the bit rate changes are high, due to more number of packets begin sent from Figure 5.20.

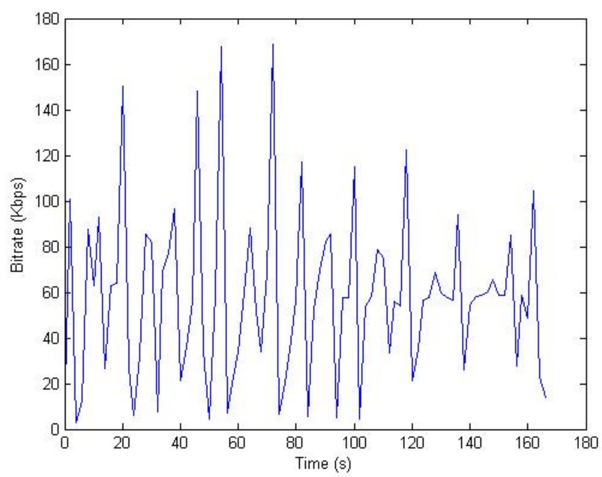


Figure 5.21. *Network utilization for a Video call on static position shows low frequency of bitrate changes over time*

As a variation to the above experiment, the mobile phone was placed in a stationary position and made to transmit a static video. As mentioned earlier, the codec used for video transmission is of variable bit rate and changes even for static pictures. Comparing

Figure 5.20 with 5.21, it can be seen that the frequency of change in bit rate is much lower in the static picture. All the measurements were taken at a probing interval of 2 seconds, *i.e.*, the service probes the file every 2 seconds for collecting utilization values. The results can be continuous if the sampling interval is small.

5.4.3. *CPU Utilization of SIPUA*

This experiment attempts to find the amount of CPU used by the process 'sipua', which is the SIP user agent for the mobile. First the cpu utilization was tested when the phone was in idle state.

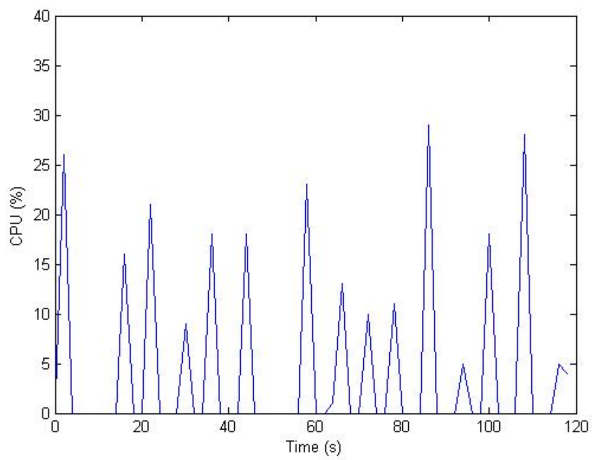


Figure 5.22. *CPU utilization when application is Idle and the spikes were due to client's keep alive mechanism*

From Figure 5.22 it can be observed that the spikes in the CPU utilization range between 0 to 25%. As discussed earlier for Figure 5.18, the SIPUA initiates a keep alive mechanism in regular intervals of time to describe its presence to the server. This timely probing of the server takes up some of the CPU resources given by the spikes in data.

The most power consuming peripherals of the mobile phone are the display, camera and the speaker phone. An experiment was conducted to identify the load over the CPU when the camera is turned ON. It was possible to obtain a resolution of a sample for every 8 seconds, even though the monitoring was set to a sampling rate of 2 seconds.

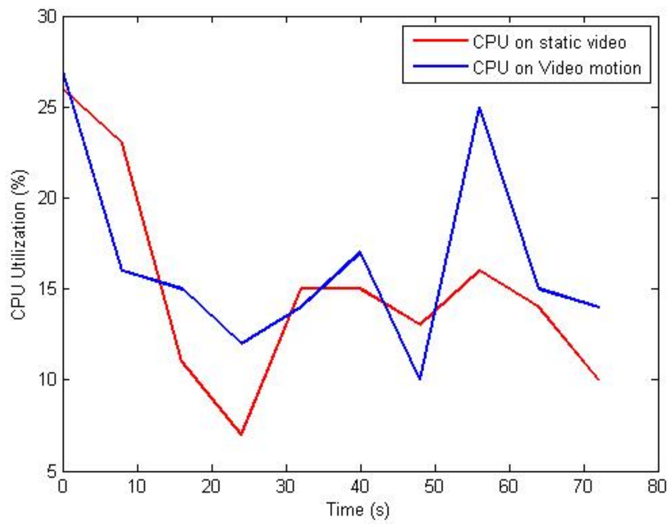


Figure 5.23. *CPU utilization when Video is turned ON: The video consumes a maximum 26 % of CPU while transmitting video at CIF(352x288) resolution*

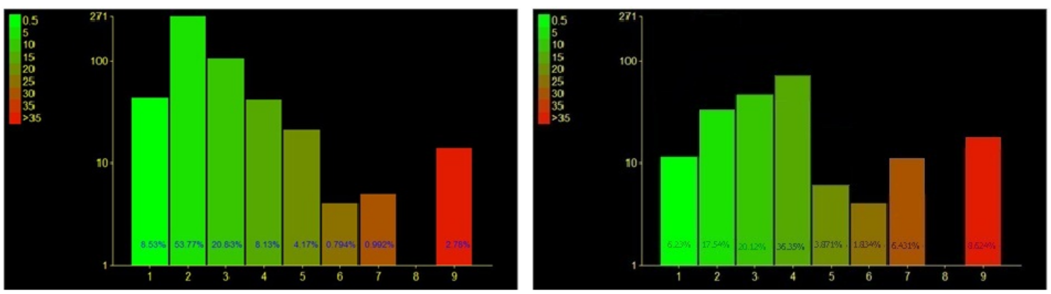
From Figure 5.23 it can be concluded that the video module takes up an average of 12-18% of CPU all the time. It initially starts with greater CPU utilization, because of setting the parameters and starting the camera, and drops down to mean level after a short while. It can also be seen that there is only a minimal impact on the CPU usage by video taken while walking and staying still.

5.4.4. *Call Quality Analysis of System*

For this experiment a commercial tool 'Hammer Call Analyzer' [38] was used to monitor the audio quality during a call. The hammer call analyzer is an application-aware analysis tool that uses a passive approach to monitor and capture network traffic. The analyzer has many features such as media analysis, VoIP decoder, multi stage call flow display, etc. The audio quality is often described by the Mean Opinion Score (MOS)values which scales from 1(bad) to 5(best). The tool predicts the human rating behavior and assigns the value for

MOS. The estimation of MOS depends on a prime component called the R-Factor. The tool follows ITU E-model [39] in calculating R-Factor which considers end-to-end delay, echoes, side-tones, loudness and other factors along with speech quality.

(i) Stream Quality Signature (SQS) shows the frequency and distribution of inter-packet arrival variation during a call. The distribution is shown over nine time bins with static ranges. The bins are color-coded and arranged on the X-axis in ascending order. The number of packets that fit each bin is shown on the Y-axis using a logarithmic scale. For each bin, the display shows the percentage of total packets contained in the bin. The percentage shown in the bin is the number of packets with an SQS score in the range shown in the legend of Figure 5.24 (a)



(a) (b)

Figure 5.24. *(a) SQS metric without Video transmission: most packets fall into 1-3 bins which states high audio quality (b) SQS metric with video transmission: The audio packet inter-arrival time increases and falls in 4-9 bins, which depicts poor quality of audio than (a)*

With greater number of packets in the first three bins, the quality of the audio is good. Figure 5.24 (a) was obtained for the audio played in a loop by a music player without video transmission with normal microphone gain. The same experiment was repeated with the video transmission to the other end and the microphone gain at maximum. The results showed the number of packets in the first three bins with video transmission was very low as in Figure 5.24 (b)when compared with 5.24 (a). Hence it can be concluded that the audio quality gets worse because of video transmission and interferes with the packet transmission in the mobile phone.

(ii) Table 5.4 shows some of the useful metrics used for calculating the R-Factor. The audio quality is determined based on the R-Factor where the value 100 being the best, 70 being minimal quality for telephony conversation and 0 begin the worst. In the case of an audio stream with normal microphone gain *i.e.*, with the speaker phones OFF, value of R-factor obtained was 81, where as the R-Factor was 62 for the audio stream with the microphone gain at maximum when the speaker phone was ON.

Table 5.4. *Comparison of Microphone quality during voice transmission with normal and Maximum gain*

Metric	Speaker Phone OFF	Speaker Phone ON
Payload Type	PCMA	PCMA
Received Packets	506	612
Lost Packets	0	1
Out of Sequence Packets	0	0
Duped Packets	0	0
Jitter	60	94
Maximum Jitter	177	242
R-Factor	81	62
Mean Opinion Score	3.875	3.045

Based on the metrics, the MOS was calculated to be 3.875 for microphone gain at normal and 3.04 at maximum. A MOS score of 5 indicates the best audio quality. Thus it can be concluded that the audio quality was good with normal gain and speaker phone turned off.

5.4.4.1. *Voice Quality under Different Microphone Modes.* The mobile phone audio operates in three different modes namely routing via the ear piece , via the speaker phone and through the hands free kit. The microphone is built in a way to accustom to these three modes for maximum performance. In the ear piece mode, the microphone gain will be low to capture the audio, spoken close to the mouth piece. When the audio is routed to the speaker phone, the gain of the microphone is increased to receive audio from signals of a certain distance. The mobile phone enters the hand free mode as soon as the kit is plugged in. The hands free kit contains a dedicated microphone and doesn't use the built-in microphone. A small experiment was conducted to find out the best mode for good audio quality. A tool called 'OmniPeek' [40] was used from Wildcard to analyze the call quality by capturing the packets on the fly. The application calculates the MOS, R-factor, delay and jitter for the audio and presents a qualitative value. For experimental data in the three modes, the results tabulated in Table 5.5 were obtained.

Table 5.5. *Audio quality under different audio modes*

Mode	Jitter (Sec)	R-factor	MOS
Ear Piece	0.063032	76	3.69
Hands Free	0.059344	85	4.00
Speaker phone	0.058206	60	2.96

Based on our result, it can be observed that the hands free mode gives the best quality, due to its elimination of most of the background noise and the echo. Also, talking in the ear piece had considerably lower quality when compared with the hand free kit. The quality of audio was the lowest in the speaker phone mode because most of the background noise was captured along with the speakers' voice. The microphone in the android was too sensitive to capture even noises from 6 to 10 feet in a quite environment. In addition to this, the speaker phone gave a feedback to the microphone, making it to echo for the receiver. On a general case, the jitter in the overall system was high, which degraded the performance of the system. Reducing the jitter and processing the noise would yield a good reception in audio quality.

5.5. Conclusion

This chapter proposed a new method for remotely controlling the mobile phone over a SIP call. The module for image transfer using SIP was developed and its performance over the network was evaluated. The call quality analysis showed that the presence of video affects the quality of a call during transmission. Similarly, the audio quality under different microphone profiles was investigated. Results showed that the hands free mode gives the best audio quality. Results of the CPU utilization studies using Nexus One, showed that the CPU utilization never exceeded 25% of the CPU load while using the camera. This indicates that there is a considerable amount of power left in the CPU. This proposed system could be integrated with the vital sign diagnostic modules for transmitting realtime physical status of a subject.

CHAPTER 6

SOCIO-TECHNICAL ASPECTS OF VIDEO PHONES

6.1. Introduction

When the telephone and e-mail are juxtaposed against each other, the telephone emerges as the leader of connectivity, supplying a real-time audible link between two people as opposed to a merely literal connection offered by e-mail. The video phone, however, is superior to both of these two forms of communication. Not only does the video phone allow an audible connection between two users like a telephone, but it also provides a live visual feed of both parties, bringing the conversation to a level, that rivals that of face-to-face communication. The conversation as virtual as this one between parties stationed as far away from each other as Japan and Munich, for example, signals a major breakthrough in the communication industry.

Despite its advantages, the video phone has not achieved widespread use. The delay for a massive deployment of such a technique is not only due to technical issues but also due to social problems. The goal of this paper is to describe some of the existing social issues and bring some quantification to technical issues; currently deployed video phones and their broader impacts on society; privacy issues related to the use of video phone technology and analysis of the relation between the perceived quality of video with respect to quantization, frame rate, and social factors like the environment of the user.

6.2. Social Impacts

The ability to read facial reaction, body language, has proven to be a much more important medium for communication between persons compared to solely audio correspondence [41]. Point-to-point video communications are deemed useful by a myriad of people. From chief executive officers to lay-persons, point-to-point video phones can give the closeness and reality of an actual face-to-face conversation needed for everyday life. For an attorney, to see clients or potential witnesses face to face and to observe their facial expressions and body language may be crucial for a cases outcome. This capability is similarly useful in conversations between doctor and patients because it allows visual analysis of the patient. Frequently, crucial business meetings taking only a matter of hours may require executives to travel thousands of miles. With video phones, virtual meetings can be held with ease thousands of miles away, thus saving both time and money.

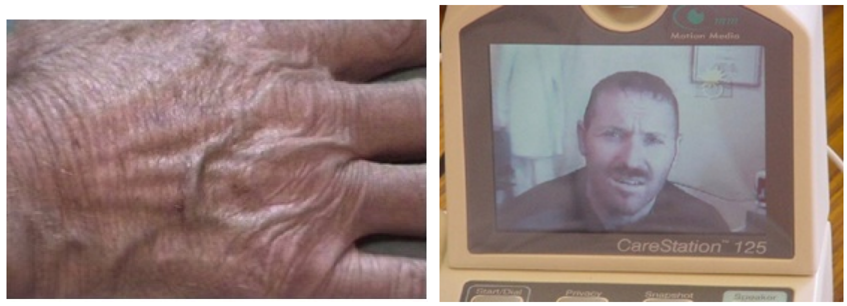
Video phone conferencing is one of the technologies that allow people to talk face to face to clients sitting in another part of the globe. This technology is not only used for businesses but also for interviews, lectures, and so on. In broader terms, video phone conferencing can be divided into two types: point-to-point conferencing and multi-point conferencing [42]. Point-to-point conferencing refers to communication directly linking two sites, whereas multi-point conferencing is communicating with three or more sites at the same time. This section describes a few of the many studies showing how the use of video phones has had a huge impact on peoples lives. Studies in rural Missouri analyzed the effectiveness of using point-to-point videophone conferencing for a 3-day professional development workshop of elementary school science teachers. The results suggest that the teachers perceived video conferencing to be as effective as traditional on site professional development workshops [43]. With conversations being much more personal and involved, the possibilities for communication are virtually endless.

Washington State uses video phones to reduce the need for public health nurses to travel to patient's homes. In Japan, medical staff uses video phones to make direct observations of inhaler use and assist with exacerbation management in Japanese patients with severe asthma. There is also a study proving the usefulness of video phones in the treatment of patients with dementia and in reminiscence therapy [44]. This use of video phones allows staff to improve asthma control and reduce hospital admissions. In Tokyo, a video phone system has been used to provide respiratory-care specialists resources to primary-care physicians and to pediatric patients requiring home ventilator support. This use resulted in large reductions in unscheduled visits by patients, home visits by physicians, and hospital admission days [45].

Elderly deaf people cannot communicate effectively, cant listen to the radio or television like normal people, and have limited access to a major part of the world. Their isolation from others is enormous. Even worse, they have to depend on others for making a telephone call, even to get to a doctor. With a video phone, a hearing-impaired person can communicate by lip reading or by using sign language either directly with the other party or with an interpreter who then translates to the other party. Those in the deaf community who use the videophone depend on it as not only a means to sign to persons fluent in sign language but also to communicate with the non-fluent [46]. One study describes Bristol City Councils joint venture with British Telecom to supply 40 houses with video phones in which 30% of the houses have elderly people with hearing impairments. This study used an interpreter service and tested its usability.

It is a commonly addressed problem that some of the needs of the elderly people who are discharged from a hospital are not satisfied. They suffer drug non-cooperation, isolation, and limitations to access specialist from their very home. In these situations, elderly people can be managed from their homes by using video phones. A medical staff can watch patients and access their records from their offices without being in a hospital ward for rehabilitation [47].

A study was made to examine the effects of tele-medicine technology on communication by comparing the style and content of communication between actual (*i.e.*, face to face) and virtual (*i.e.*, non-face to face, videophone) dermatology visits. The hypothesis was that there is no difference in the content and style of communication between actual and virtual visits in dermatology. [48]



(a) (b)

Figure 6.1. *(a) Inspection of symptoms (b) View of a patient under diagnosis*

Its commonly seen that remote places on a map do not have hospitals, but it does nt mean that people do not live there or that they dont get sick. Remote places do not have hospitals because its too expensive to get all the required medical facilities to the rural areas. Hence, the term virtual clinic was coined. Virtual clinics use low-cost video phones (Figure 6.1 (b)), together with some of the clinical equipment mediated by a trained local volunteer to provide remote consultation and monitoring of patient records. This system potentially saves long trips into town by patients since the traditional home visit is not feasible in these regions. A virtual clinic was set up in the rural region of Victoria, Australia by the Department of Rural Health to avoid some of the chronic fatalities due to lack of basic medical services and difficulty for elderly people to reach the medical facilities in time [49].

Figure 6.2. *Video Conferencing with grandchildren*

Video phones may also be deployed in houses for personal use, where each person can talk as well as see the other person, however remotely located they may be in the actual world.

A grandmother could be talking from her home in London, while her son and family could be responding from their home in New York. She could actually see her newborn grandson from across the Atlantic without traveling thousands of miles (Figure 6.2). People can have the peace of mind to check their baby and babysitter from anywhere in the world when they are away. Seeing the person with own eyes and talking to them gives immense satisfaction and creates the trust that they are talking to the right person and everything is under control.

Video phones can allow people to record video, images of their loved ones, family members, and special occasions. They can also be used by soldiers in the battlefield to talk face to face with their families in the United States, allowing them to voice their emotions. With all this huge impact on human lives, its not that hard to use a video phone. All that is needed is an Internet connection with a voice over Internet protocol (VoIP) service to connect with almost anyone in the world at the lowest cost.

6.3. Social Issues

The degree to which the problems caused by video phones affect society may seem less serious than those of other misused technologies. Nevertheless, the issues concerning video phones can be ignored before they may become a part of our daily life. Currently, video phone screens require both speakers to be looking at the person they are talking with as well as seeing them on the screen. Even though it may seem so, the screens focal point is not the same as the cameras focal point; hence, eye-to-eye contact is not possible. Nevertheless, inflections, expressions, and other non-verbal features lost in cyber-space can be preserved in video telephony, helping to reconnect people during lifes special moments [41]. However, certain social issues still prevent wide-scale deployments. Thus, as video phones poise to take off, there is a need for a closer look at the possible problems of using video phones on a daily basis.

6.3.1. *Display of Mood and Emotions*

Video phones can cause additional loss of privacy due to awkward situations while answering a video call. When using a videophone, however, both parties may become uncomfortable if the person they see has not dressed in a manner they consider appropriate or professional (Figure 6.3). Video phone calls, like phone calls, intrude on a persons privacy. Therefore, any video phone system of the future must balance retaining privacy to responding to emergency communications before wide-scale deployment of video phones will be successful.

Figure 6.3. *The called person may or ny not be presentable to the caller*

6.3.2. *Trustworthiness*

Video phones cannot develop similar trust as much as in face-to-face conversations. Just as phishing occurs in e-mail, it is expected that it will soon find a way into video phones. As with Internet phishing, criminals may use the pretext of verification of trivial details or transactions. With a look-alike backdrop and realistic acting, these criminals may convince users that they are authorized to request confidential details, such as code/PIN numbers. Currently, video phone users have no definite way to distinguish among the people purportedly calling from a bank or to authenticate whether callers are really associated with the firm they claim to represent.

6.3.3. *Unexpected Video Clips*

Video mail, an application similar to voice mail, might come into use as video phone technology expands. However, video spammers can exploit video mail by flooding video-mail inbox with unwanted videos. These video mails could be important messages and family issues. They also could be video spam; furthermore, unwanted video clips or messages that may be distributed over the network may be inappropriate to view.

6.3.4. *No Control over Surroundings*

Within certain situations, a mobile video phone can make communication easier, but in other situations, it can become a nuisance. Foremost among such nuisances is loss of privacy. As with standard cell phones, the user could be anywhere when the mobile video phone rings. Attending a call on a mobile phone does not make the caller worry about his visual appearance or his current activity [50], [51]. A mobile video would give the user little choice whether he would like to be seen or avoid being seen and disclose him to outsiders [50], [51].

6.4. Technical Issues

In addition to the social issues described in section 6.3, there are a number of technical and security-related issues that demand fast and secure solution before the widespread use of video phones. Some of these issues are identified and described in this section.

6.4.1. *PC-Based Video Conferencing*

Currently, the most popular form of video conferencing is via the PC in which the user needs a computer, the camera, and optionally headphones in order to communicate via video with the receiver. Video conferencing has assisted communication in both the corporate and the civilian world with moderately user-friendly software structurally comparable to instant messaging. Video conferencing has also shown potential in the medical world, making the formerly defunct idea traditional house-calls feasible again. Although PC-based video conferencing and video phones both support point-to-point and point-to-multi point video communications, considerable technical differences separate these two devices, as described in Table 6.1.

Table 6.1. *PC Video Conferencing vs Video Phone*

PC-Based Video Conference	Video Phones
Requires a centralized video server	VoIP infrastructure can be reused
Requires complex configuration, lacks continuous availability, and requires maintenance	Plug-and-play device
Inherits all the PC-based security issues	Dedicated operating system and hardware; hence, comparatively fewer vulnerabilities
High-performance hardware; hence, scalable to a large number of participants in a multi-point video conference. In addition, can hold large amount of video mail	Low cost hardware; hence cannot support large size multi-point video conference; small number of video mails due to limited storage space
Lacks inter-operability between different conferencing systems	Proven inter-operability due to already existing VoIP equipment
Complex user interfaces; can be an issue for wide-scale deployment across the masses	Easy to use and operate by residential users

6.4.2. *Comparison of Social Issues in Three Electronic Communication Systems*

Any kind of communication system has some vulnerabilities and privacy issues with respect to user context. Table 6.2 compares some the common privacy and social issues for the most popular means of communication.

6.4.3. *Size of the Video Screen*

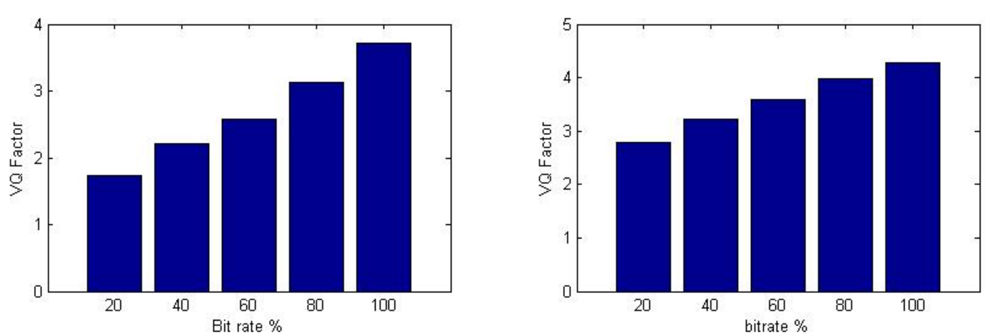
Video conferencing requires adequate video quality to be able to simulate a face-to-face conversation. In many situations, only a limited bandwidth is available (wide-area network), so maximizing its efficiency is crucial. Several video codecs are available now to satisfy the

Table 6.2. *Social Issues of Electronic Communications*

Issue	E-Mail	Voice	Video
Unwanted Calls	Annoying	Annoying, possibly embarrassing	Annoying, embarrassing, and potentially harmful
Phishing	People generally ignore e-mail from unknown source	People follow up on voice calls with other reality checks	High chances of impersonation and easy spoofing
Privacy	Bystanders cannot access without permission	Bystanders overhear conversations without permission; but callers can hide mood and emotions	Physical surroundings, mood, emotions, and all the details of the callee are disclosed
Junk Mail	Impacts productivity	Nuisance	Embarrassing; viewer discretion required
Presence	Difficult to find the location	Special location services required	Location and mood automatically disclosed

needs for quality. Among all the services, teledata [52] requires the most precise details, which means that communication actively involves video rather than audio. Some of the services require higher frame rates to provide a harmonious and synchronous conversation. For example, there can be visual content during the course of a lecture where the frame rate of about 15 to 25 per second will be considered adequate to meet the need. In some applications [53], even 2 frames per second is enough. In another context, consider hearing-impaired people who are trying to have a conversation. A frame rate of a minimum of 25 frames per second will allow lip reading, and audio does not have any influence in this case. So a specific level of requirement is needed for a particular group of users. Choosing the optimal requirement is another significant measure to keep the network distant from congestion.

Some of the most popularly used codecs in video conferencing are H.263, H.264, H.261, and MPEG standards, in which the bit rate is typically n x 64 kbps. One frequently asked question is what image resolution (and subsequently, screen size) is best for social interactions, emotions, and expressions. Experiments were carried out to find the factors that could considerably affect video quality. In Figure 6.4 (a), a measure of the quality of the video transmitted using real-time transport protocol (RTP) is made against the bit rate; under a common intermediate format (352 x 288) video using H.261 (ITU-T video coding standard) codec. Figure 6.4 (a) gives the quality of the picture, quantized over a metric VQ factor (vector quantization factor) [53] on a scale of 0 to 5, 5 being the best video clarity and 0 being the worst.



(a) (b)

Figure 6.4. *(a) VQ factor vs bit rate (varied from 20% to 100% of the available bandwidth, CIF format) (b) VQ factor vs bit rate (varied from 20% to 100% of the available bandwidth, QCIF format)*

The VQ factor is calculated by using an algorithm which is proprietary to the software ClearSight [54]. With an increase in the number of bits per second, it can be seen that the quality of received pictures getting better but at the cost of bandwidth. In Figure 6.4 (b), the quality is measured for a video using QCIF format (176 x 144), with measurements of VQ factor over bit rate. From Figure 6.4 (b), it can also be seen that the quality of the video does not suffer much with the increase of the video resolution because the image size is directly related to the bandwidth it occupies in the network (e.g., a full-size video conferencing on PC compared to small screen video phone).

6.4.4. *Perception of Motion and Distance*

One of the vital necessities for the video phone is to understand the emotional quotient of the person during a conversation. Video stream can incur a loss in the clarity of the picture due to encoding scheme and the network loss. Furthermore, some of the factors of the surroundings, like the users distance from the camera and brightness, account for the correct interpretation of actions at the other end. Vendors will frequently update the codec on the video phone and increase the price. So, another frequently asked question is what codec is optimum for social interactions. To address this issue, measurements have been done to determine the mean opinion scores (MOSs) based on motion, distance of the subject from the camera, and surroundings of the scene.

6.4.4.1. *Distance from the Camera.* In video phone usage, the user can be too close to the phone or several feet away from the phone. Whenever the user performs an action or makes a movement, it can be identified by the number of pixel changes in the video frame. These changes in the pixels are measured by the brightness flickering metric (BFM) [55], which gives the amplitude of the pixel change from the previous frame.

Experiments were conducted using two different video encoding schemes, H263 and H264, using the soft phone X-lite. The measurements are made by using the video phone at different

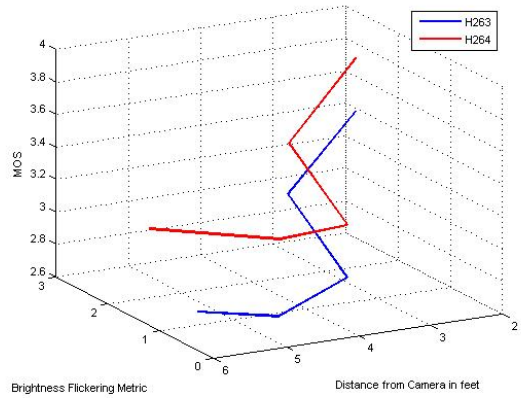
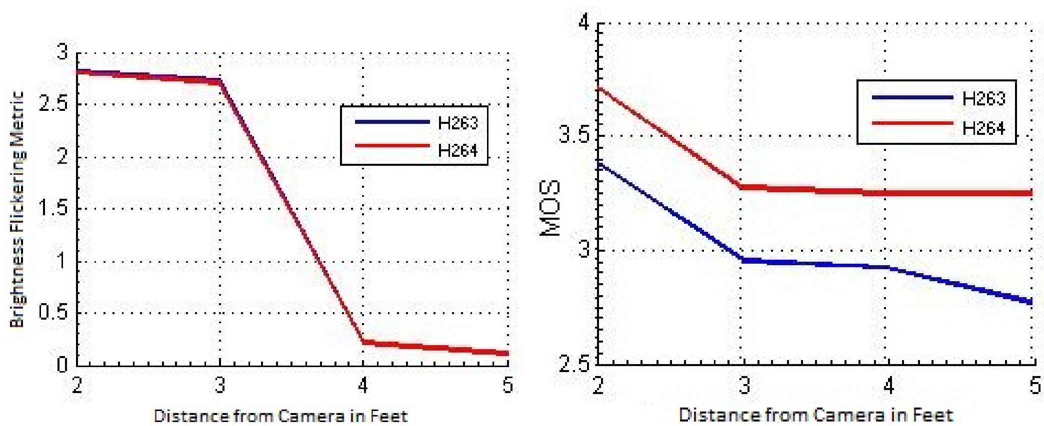


Figure 6.5. *Motion detection with respect to distance from the camera*

distances from a single user (to avoid subject-specific measurement errors). The subject (user) is subjected to a uniform activity for all the measurements taken, also the bit rate is set to 128 kbps for optimum performance. As seen in Figure 6.5 and 6.6 (a), the quality of



(a) (b)

Figure 6.6. *(a) perceived quality vs distance from the camera (b) Brightness flickering metric Vs distance*

the video given by the MOS for the video format H264 is better than the quality provided by H263. In addition, the quality of video gradually deteriorates with the increase in the distance between the phone and the subject. Thus, it can be inferred that the quality of the video can decrease only up to a certain distance and that it will continue to remain the same or increase as the pixel changes of the frame become negligible. From Figure 6.6 (b), it was inferred that the detection of motion from the video decreases with increasing distance from the phone and that the perceived quality of video has the least impact on the perceived quality of the video.

6.4.4.2. *Brightness of the Environment.* With the widespread capabilities and the ability to start a face-to-face conversation from almost anywhere in the world, it is expected for the videophones to be used almost everywhere, from parks and beaches to mines and tunnels. Owing to expanse of its usage, it is obvious to find out whether the videophone can meet its primary objective of delivering facial reactions in different brightness and light sources. A human eye can adapt to the different brightness level with ease, but a phone camera cannot do so as easily. Even with the advent to new technologies for aperture control and image stabilizations in recent video phones, the quality of the captured video greatly differs from that of natural human vision. To identify how the brightness of the surrounding affects the

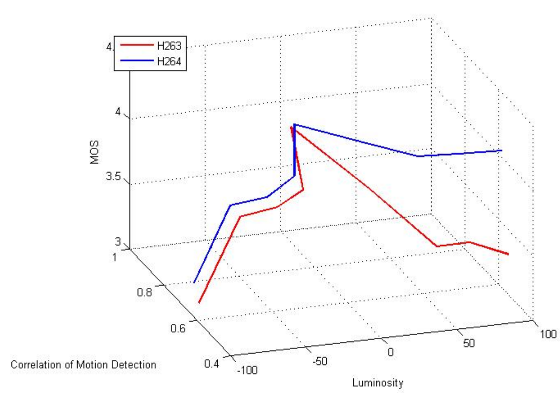
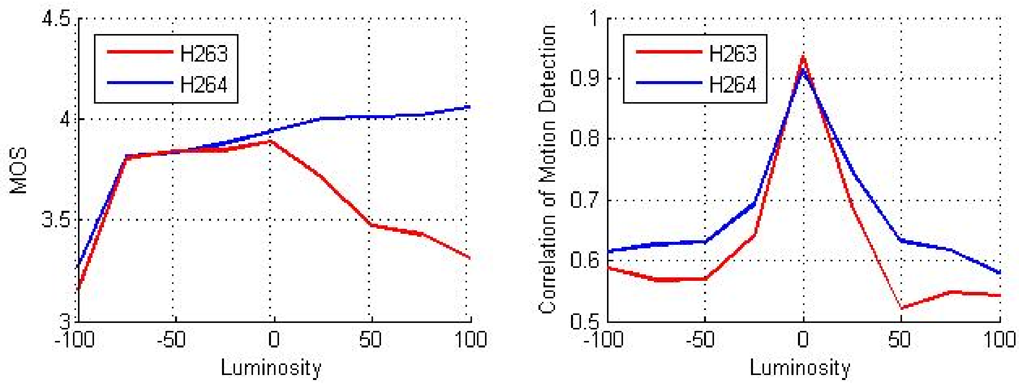


Figure 6.7. *Motion detection with respect to brightness*

ability to detect motion changes in a video phone, real-life scenes were used with different levels of brightness from very high to too low and made the test subject (user) to replicate the same kind of motion in all the scenarios. Preliminary experiments were made to identify the optimum lighting conditions for motion detection, which is taken as zero luminosity or reference. From the results in Figure 6.7 and 6.8 (a), it was observed that the encoding



(a) (b)

Figure 6.8. *(a) Quality of video vs Brightness (b) Motion detection Vs Brightness*

scheme H264 performs better in brighter lighting conditions and that the perceived quality of both the codecs H263 and H264 are the same in dim or poor lighting conditions. When considering the H263 video codec, the quality observed was greater in normal and medium-low conditions than the other extremities. H264 encoding scheme occupies more bandwidth to retain the frames under higher lighting conditions giving increased quality compared to H263. A measure taken against the luminosity and the motion detection using the brightness flickering metric shows gradual decreases as the values was moved away from the normal in both directions from the results in Figure 6.8 (b)

The visual quality of the video is mainly considered by the blockiness [55] of the received picture and the amount of pixilation in the frame. The quality of the video drops down drastically with high levels of blockiness in the frame. Based on the test results from Figure 6.9, it can be inferred that the blockiness of the video increases dramatically in the presence of brighter light than the poor ones. It is observed that whenever there is a motion or movement in a region of higher spectral intensity, the video capture results in the production of blocks of

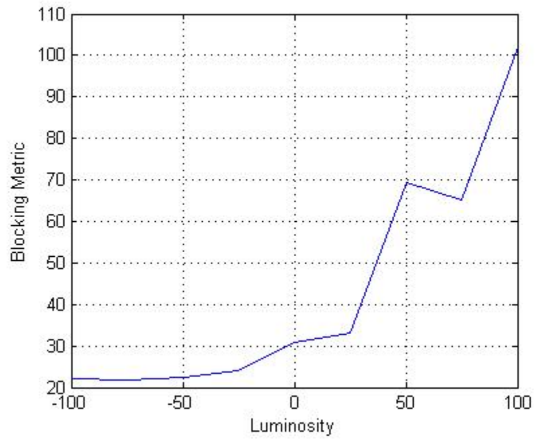


Figure 6.9. *Blocking Vs Brightness*

pixels around the object in motion. As the light gets darker, the differentiation between most pixels was hard; *i.e.*, the blocking lowers and eventually gets to zero, tending to complete blankness.

6.4.5. *Change in Bandwidth Due to Human Activities*

The frame rate for the whole experiment was set at 24 FPS (frames per second). Hard phones were used as well as soft phones for the measurements. The rates obtained for different resolutions (first three are for hard phone and last two are for soft phone) are tabulated. From Table 6.3, it can be seen that the bit rate increases as the resolution of the captured image increases. Also it can be observed that every human activity has a direct impact on the bit rate associated with the frame in the video stream. For example, in the absence of a user in the view port indicating a completely static scene, the bit rate maintains a very low value of 4Kbps for every resolution. However, there is a change in bit rate when the scene includes person even he is in complete stand still.

The increase in the bit rate for different activities for the same resolution is due to the movement of exposed body parts involved in the activity. Hence it can be concluded that bit rate increases even with the small amount of motion exerted by the human body without the user's knowledge. These changes are evident when person is breathing normally and heavily.

Table 6.3. *Bit rate for different human activities in Kbps. First three columns are collected from hard phones whereas last three are from soft phone.*

Activity	176x144 (QCIF)	352x288 (CIF)	704x576 (4CIF)	320x240 (QVGA)	640x480 (VGA)
No user	4	4	4	5	5
Eye Blink	10	25	35	22	34
Smile/Scream	8	21	33	16	32
No Breathing	8	19	21	11	20
Normal Breathing	12	35	38	17	39
Heavy Breathing	20	62	75	24	71

6.5. Conclusion

In this age of e-mail and instant and text messaging, video telephony simulates the personal nuances that come from experiencing face-to-face communications [41]. The rapid growth in broadband networks has contributed to the anticipation of a similar growth in the field of video over Internet protocol; but although the video phone has been available for decades, widespread deployment has not actually happened. This lack of deployment has attributed to the need of technological improvement to support video over Internet. Certain aspects have yet to develop to trigger growth of a video phone market. Apart from the technological needs, interesting social aspects must be analyzed before video phones can become as ubiquitous as the plain old telephone system (POTS). It can be argued that PC-based video conferencing is a cheaper solution, but the issues discussed in this paper are germane to PC-based video communications. As can be seen for Figures 6.4 (a) and 6.4 (b), the small frame rate with low bandwidth is good enough and can still result in satisfactory video quality. It was also observed that H263 performs well for all the day-to-day social networking activities. Standing 4 feet from the camera can still give reasonably good video quality in the currently available codecs. However, H264 fares better in poor lighting conditions. By addressing social issues and aided by the developments in technology, fully functional video phones can become a reality and can probably pave way for a better understanding between countries and cultures.

CHAPTER 7

CONCLUSION

Remote media control is a system designed for VoIP communication between a mobile and PC client over a SIP call. All the multimedia components on the mobile are controlled from the PC client(call taker). Communication during emergency calls includes around 34 different protocols [56] where the caller is put forth a series of questions by the call-taker, to decide the kind of service to be offered for the victim. The protocol contains some time/life critical events such as cardiac arrest, fainting, etc. which need immediate attention. Also in case of traumatic events, remote media control will be useful as the call taker can take over control of the phone to visualize and adjust the mobile phone for optimal settings, thus reducing stress on the caller. Research is begin carried over to redefine the medical protocols for efficient and faster response to situations with remote media control.

Measuring the vital signs such as blood pressure, heart rate and respiratory rate of a person during the above mentioned situations could be a difficult task due to the lukewarm usage of portable devices like blood pressure and heart rate monitors. Using a mobile phone for such tasks could be a very convenient alternative. With the sensor rich mobile phones in today's market, medical applications like blood pressure and heart rate monitoring-the two most important information about a person are feasible. For instance, the camera could be used for pulse measurement, and the accelerometer for capturing respiratory patterns.

Using the capabilities of the sensors mentioned in the previous paragraph, this thesis work developed two systems, 1) Measuring vital signs of a person using the mobile phone and 2) VoIP communication for remotely controlling the mobile phone. The Cuff-based and Cuffless estimation of blood pressure was designed which resulted in average accuracies of 70% and

65% respectively. A further improvement in accuracy was obtained by the cuffless differential method which used vascular transit time. By utilizing the camera with flash, the heart rate was measured using the pulse from the capillaries in the finger. An accuracy of 87% with a 5 sec data to 99 % with a 30 second data was obtained without any calibration. By calibrating the data, an average accuracy of 92% was obtained even for 5 sec data. A technique for calculating the respiratory rate of a person through his respiratory patterns obtained using an accelerometer on the phone was explained.

The remote media control was developed with an image transfer module for enabling privacy protection to a caller incase the events are inappropriate for a video feed. The usage of the system during emergency situations was evaluated. The call quality analysis showed that the call quality was affected by the presence of video during transmission. Similarly, the audio quality under different microphone profiles was investigated. Results showed that the hands free mode gave the best audio quality. Results of the CPU utilization studies using Nexus One, showed that the CPU utilization never exceeded 25% of the CPU load while using the camera. This indicates that there is a considerable amount of power left in the CPU. This proposed system could be integrated with the vital sign diagnostic modules for transmitting realtime physical status of a subject. However, the overall evaluation of this system is slated for a future work.

7.1. Challenges and Limitations

I encountered a number of challenges in this work. In Chapter 2, the first challenge was the improper holding of the phone while measuring the heart rate. If the finger was not placed properly with the right amount of pressure, the camera did not detect the pulse from the finger. Hence the experiments with new subjects had to be repeated a number of times for accurate data. Although the calibration method resulted in a good accuracy, it is a novice method. In Chapter 3, while measuring the respiratory rate, the accelerometer did not record any changes in acceleration when the person was in a upright position. Hence a lying down position had to always be maintained or the phone had to be strapped to the chest. Also a mild breathing or sudden disturbances during normal breathing like speaking, partial breathing, etc., in a person did not yield a proper respiratory pattern. For blood pressure measurement explained in Chapter 4, the main requirement is the availability of two phones and a camera flash in atleast one of the phones. Another important necessity for the measurement is one of the phones has to be placed on a bare chest for heart sounds. Hence some users might find this method inconvenient under certain situations. Moreover, the placement of the phone is very crucial for heart sound detection. The camera's frame rate had to be higher for accurate measurements. The phone did not allow me to increase the frame rate without increasing the resolution. Hence the file size increased which led to larger processing time. Finally in Chapter 5, during a video conversation, the video lagged the audio. This caused a delay in the operations being executed after receiving an instruction. These challenges have to be addressed by developing new methodologies and operational protocols.

7.2. Future Work

In this section, I will mention some of the future work related to this thesis. Firstly, I will implement the pulse detection in real time with the ability of the mobile phone to identify and instruct a proper way of holding it. Secondly, a real time filtering of audio heart beats capable of detecting beats even without direct skin contact to the mobile phone over the chest. Thirdly, as mentioned in Chapter 3, there needs to be a better understanding and analysis of respiratory pattern through oxygen saturation as it could help in estimation of oxygen in blood flow. A method has to be developed to calculate the lung capacity using the respiratory pattern. The methodology for remote media control has to be optimized for audio and video conversations. A robust system has to be designed for improving efficiency even in the presence of noise. A new method has to developed for efficient transmission of vital signs to the dispatcher. A new set of protocols has to be defined for human machine interfacing that could effectively reduce the conversation time during an emergency call.

REFERENCES

[1] N. Abhijit, "Normal pulse rate", http://www.buzzle.com/articles/normal-pulse-rate.html.

[2] T. Gerard and A. Nicholas, *Principles of Anatomy and Physiology*, Harper-Collins, New York, sixth edition, 1990.

[3] Wikipedia, "Variation of vital signs with age", http://en.wikipedia.org/wiki/Vital_signs.

[4] S. Kirk, "Photoplethysmography: Beyond the Calculation of Arterial Oxygen Saturation and Heart Rate", *Anesthesia & Analgesia*, vol. 105, no. 6S Suppl, pp. S31–S36, December 2007.

[5] Google Inc., "Nexus one: Technical specifications", http://www.google.com/phone/static/en_US-nexusone_tech_specs.html.

[6] K. Banitsas, P. Pelegris, T. Orbach, D. Cavouras, K. Sidiropoulos, and S. Kostopoulos, "A simple algorithm to monitor hr for real time treatment applications", in *Information Technology and Applications in Biomedicine, 2009. ITAB 2009. 9th International Conference on*, 4-7 2009, pp. 1 –5.

[7] "Medicine & science in sports & exercise", *The Official Journal of the American College of Sports Medicine*, vol. 39, no. 5, pp. 749–898, 2007.

[8] G. Ronald, G. Brian, O. Ronald, M. Audry, R. Gary, and M. Virinder, "Longitudinal modeling of the relationship between age and maximal heart rate", *Medicine & Science in Sports & Exercise*, vol. 39, no. 5, pp. 503–508, 2007.

[9] "What should you know about your heart rate or pulse", http://www.nemahealth.org/programs/healthcare/heart_rate_pulse.htm.

[10] W. Kenneth, H. Dallas, and H. Willis, *Clinical Methods, The History, Physical, and Laboratory Examinations*, Butterworth-Heinemann, Boston, third edition, 1990.

[11] G. Mary, "Pulse oximetry", *Crit Care Nurse*, vol. 22, no. 3, pp. 69–74, 2002.

[12] P. Leonard, T. Beattie, P. Addison, and J. Watson, "Standard pulse oximeters can be used to monitor respiratory rate", *Emergency medicine journal : EMJ*, vol. 20, no. 6, pp. 524–525, 2003.

[13] A. Donald, S. David, D. Renee', and H. James, "Estimation of airway obstruction using oximeter plethysmograph waveform data", *Respiratory Research*, vol. 6, no. 1, pp. 65, 2005.

[14] B. Martin, "Aacn practice alert: Non-invasive blood pressure monitoring", http://www.aacn.org/WD/Practice/Docs/PracticeAlerts/NIBP Monitoring 04-2010 final.pdf.

[15] J. Jhalani, T. Goyal, L. Clemow, J. Schwartz, T. Pickering, and W. Gerin, "Anxiety and outcome expectations predict the white-coat effect", *Lippincott Williams & Wilkins, Inc*, pp. 317–319, December 2005.

[16] "Automatic wrist blood pressure monitor", http://www.homedics.com/home/wellness/blood-pressure-monitors/automatic-wrist-blood-pressure-monitor-4.html.

[17] M. Zamir, *The Physics of Conronary Blood Flow*, Springer, New York, first edition, 2005.

[18] G. de Simone, M. Roman, M. Koren, G. Mensah, A. Ganau, and R. Devereux, "Stroke volume/pulse pressure ratio and cardiovascular risk in arterial hypertension", *Hypertension*, vol. 33, pp. 800–805, March 1999.

[19] C. Jay, F. Stanley, M. Gary, M. Dennis, L. Lisa, R. Jennifer, and M. James, "Noninvasive pulse wave analysis for the early detection of vascular disease", *Hypertension*, vol. 26, no. 3, pp. 503–508, 1995.

[20] J. Nürnberger, A. Opazo Saez, S. Dammer, A. Mitchell, R. Wenzel, T. Philipp, and R. Schaumlfers, "Left ventricular ejection time: a potential determinant of pulse wave velocity in young, healthy males", *J Hypertens*, vol. 21, no. 11, pp. 2125–32, 2003.

[21] A. Jose, W. Gabriel, G. Carlos, and C. Mario, "Contribution of stroke volume to the change in pulse pressure pattern with age", *Hypertension*, vol. 34, no. 4, pp. 808–812, 1999.

[22] D. James, "Human physiology: Circulatory system", http://www.biosbcc.net /doohan/sample/htm/COandMAPhtm.htm.

[23] T. William, *The Research Methods Knowledge Base*, Atomic Dog, Cincinnati, OH, second edition, 2000.

[24] S. David, "T- test table", http://changingminds.org/explanations/research/analysis/t-test_table.htm.

[25] B. Lynn, *Bate's Guide to Physical Examination and History Taking*, Lippincott Williams & Wilkins, eighth edition, 2002.

[26] "Heart sounds types and locations", http://en.wikipedia.org/wiki/Heart_sounds.

[27] HTC Corporation, "Htc hero: Technical specifications", http://www.htc.com/www /product/hero/specification.html.

[28] J. Foo, C. Lim, and P. Wang, "Evaluation of blood pressure changes using vascular transit time", *Physiological Measurement*, vol. 27, no. 8, 2006.

[29] IEEE 1588-2008, "Standard for a precision clock synchronization protocol for networked measurement and control systems", http://ieee1588.nist.gov/PTTI_draft_final.pdf.

[30] "Bluez: Bluetooth stack for android", http://source.android.com/porting/bluetooth.html.

[31] i-p-tel GmbH, "Sipdroid", http://code.google.com/p/sipdroid/.

[32] V. Luca, "Mjsip-mini-tutorial", http://moya.iyard.org/pub/Technology/JavaSIP/mjsip_minitutorial_01.pdf, 2005.

[33] S. Kevin, R. Neil, and L. Yves, *MBONE: Multicasting Tomorrow's Internet*, John Wiley & Sons Inc, first edition, 1996.

[34] "Session initiation protocol (SIP) extension for instant messaging", RFC 3428, Internet Engineering Task Force, Dec. 2002.

[35] J. Rosenberg, S. Henning, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: session initiation protocol", RFC 3261, Internet Engineering Task Force, June 2002.

[36] A. Roach C. Jennings, R. Mahy, "Relay extensions for the message session relay protocol (msrp)", RFC 4976, Internet Engineering Task Force, Sept. 2007.

[37] "Jpg quality comparison chart", http://en.wikipedia.org/wiki/Wikipedia:Featured_picture_candidates/Image:Quality_comaprison_jpg_vs_saveforweb.jpg.

[38] "Hammer call analyzer", http://www.empirix.com/products/hammer_call_analyzer.asp.

[39] J. Bergstra and C. Middelburg, "Itu-t recommendation g.107 : The e-model, a computational model for use in transmission planning", Tech. Rep., 2003.

[40] "Omnipeek: Network analyzer", http://www.wildpackets.com/products/network_analysis_and_monitoring/omnipeek_network_analyzer/voip_monitoring.

[41] K. Brent, "Wainhouse research: Video telephony opportunities and offerings in the new video age", http://www.wrplatinum.com/Downloads/4574.aspx.

[42] D. Andrew and W. Ira, "The business case for videoconferencing", Wainhouse Research.

[43] A. Leonard and D. Daniel, "Integrating point-to-point videoconferencing into professional development of rural elementary school science teachers", *Contemporary Issues in Technology and Teacher Education*, vol. 6, no. 4, pp. 399–413, December 2006.

[44] K. Noriaki, A. Shinji, Y. Kiyoshi, and K. Kazuhiro, "Networked reminiscence therapy for individuals with dementia by using photo and video sharing", in *Assets '06: Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*, New York, NY, USA, 2006, pp. 125–132, ACM.

[45] R. Gerard, "Telemedicine and the pediatric tertiary care center", *Ochsner Journal*, vol. 5, no. 2, pp. 11–14, 2003.

[46] S. Joyce, "Employing assistive technology in the workplace", *Hearing Health Magazine*, vol. 20, no. 4, 2004.

[47] D. George, S. Cheryl, and S. Laura, "An evaluation framework for a rural home-based telerehabilitation network", *J. Med. Syst.*, vol. 29, no. 6, pp. 595–603, 2005.

[48] D. George, E. Karen, and V. Santosh, "A comparison of communication models of traditional and video-mediated health care delivery", *International Journal of Medical Informatics*, vol. 74, no. 10, pp. 851 – 856, 2005, Supporting Communication in Health Care.

[49] B. Meade and J. Dunbar, "A virtual clinic: telemetric assessment and monitoring for rural and remote areas", *Rural and Remote Health 4*, , no. 296, 2004.

[50] V. Antonio, C. Gabriella, and C. Fabrizio, "A new upnp architecture for distributed video voice over ip", in *MUM '06: Proceedings of the 5th international conference on Mobile and ubiquitous multimedia*, New York, NY, USA, 2006, p. 2, ACM.

[51] L. Frank and P. Basavaraj, "Addressing security challenges of mobility and firewalls in voip networks", in *IEEE GLOBECOM'2004: First Workshop on VoIP Security*, 2004.

[52] A. Anne, M. Jim, J. Matthew, and S. Lucy, "Audio and video guidelines for networked multimedia applications: Applying the etna taxonomy", EPSRC ETNA Project.

[53] M. John, M. Angela Sasse, and M. Dimitrios, "Sharp or smooth? comparing the effects of quantization vs. frame rate for streamed video", in *In CHI 04: Proceedings of the SIGCHI conference on Human factors in computing systems*. 2004, pp. 535–542.

[54] "Clearsight network analyzer", http://www.clearsightnet.com.

[55] "Msu video group", http://compression.ru/video/quality_measure/index_en.html.

[56] State of New Jersey Department of Health and Senior Services, "Emergency medical dispatch guidecards", Tech. Rep., Office of Emergency Medical Services, 2006.