CHAPTER 1

INTRODUCTION

1.1 GENERAL INFORMATION

Trendy and all inclusive handheld devices are becoming dominant in telecommunication industry. The rapidity of gadgets that deliver many functions such as Video/Image processing services and associated amenities like video conferencing, video ring back, video mail and interactive-voice all in handheld mobile devices (White paper 2009). The growth also enables the manufacturer to embed digital cameras and other hardware in high end mobile handsets to support non-core services in addition to voice calls. This integration of multiple hardware units in a single device has not only eliminates the use of multiple devices but also reduces the cost while providing good quality of service. Value Added Service (VAS) is a term used in telecommunications industry for noncore services which are beyond standard voice calls. At a conceptual level, VAS adds value to the subscriber to use the mobile handsets more frequently which in-turn increases their Annual Revenue Per User (ARPU) for the service provider.

With respect to service provider perspective, value added services increases the video and image processing on a mobile phone which is benefit to both customers and also to service providers. To meet the requirements of the customer, there is a need to use efficient video and image processing algorithms. The increase in the video
processing services demands higher bandwidth and hence there is a need to use advanced pre-processing and post-processing algorithms. To implement the Video/Image processing algorithm in mobile handsets, it is necessary to reduce the complexity involved. Complexity is measured in terms of memory requirement and computational capability. Due to memory limitation in mobile hand set and limitation of bandwidth in communication channels, video coding methods become more essential and hence standardization of the video compression techniques is required. H.264/Advanced Video Coding (AVC) is open standard for Audio/Video applications which can support various applications like video broadcasting, video streaming and video conferencing over fixed & wireless networks over different transport protocols. To bring these video processing applications on a mobile phone, it is necessary to use efficient codecs such as H.264. Also some of the Video/Image processing applications can be embed on the mobile phone to increase the usage of mobile phone which is benefit to both consumer and service providers. Appendix A provides additional details about the introduction chapter.

1.2 PROBLEM STATEMENT

According to CISCO Visual Networking Index (VNI) survey report, internet video consumption has increased significantly from an average of 700 terabytes/day in 2006 to 1200 terabytes/day in 2007. The video consumption has already expanded to 650% by 2011 to 7800 terabytes/day. Video content creation will expand even more rapidly in future days. According to the same market research firm, Internet video uploads grown from 500K uploads/day in 2007 to 4800K uploads/day in 2011. For the internet video uploads on the mobile phone, bandwidth is a critical and important factor. The bandwidth can be reduced by using advanced video coding (AVC) H.264 standard.

Advanced video codec standard H.264 facilitate improved video compression by providing high video quality at significantly lower bit rates which is necessary
requirement for mobile applications. Standardized codecs also offer greater flexibility as they are able to work in a wide range of network environments. The University of Essex estimated that there is a continuous 7% year-over-year gain in compression efficiency in a paper titled “Future Performance of Video Codecs.” To measure the complexity involved in Video/Image processing applications, it is necessary to port the video clips in some encoded format on a hardware platform. Since the Video/Image processing algorithms are required for mobile applications, porting of algorithms on hardware platform is necessary. The H.264 encoded frames are decoded using H.264 decoder by porting the encoded frames on a hardware platform.

The research work presents some of Video/Image processing algorithms required for mobile applications. These algorithms on the mobile phones aims to reduce the computational complexity involved for computation and the memory requirements. The algorithms discussed can also be applied to other applications. The algorithms discussed in the research work is implemented, tested and optimized for mobile devices. The different algorithms considered are segmentation, compression, scaling and edge detection.

To be able to bring high-definition video to the hands of users, mobile device designers must optimize the power consumption of all components. While the display and wireless transmitter consume a large chunk of the total power, every component of a mobile system must be evaluated for its contribution to the power budget. Application and baseband processors must meet stringent power requirements to win a socket on a mobile handset's circuit board. Video coding is usually the most power-hungry application a mobile processor has to run. Therefore, minimizing the maximum power consumption figure of a chip often concentrates on finding the most power-efficient way to implement video processing algorithms.

The video to be transmitted over a network is encoded using H.264 advanced video coding standard. The research work focuses on decoding the encoded video and summarizing the important selected frames. The spatial redundancy of the selected
frames is minimized by using image transform coding. The compressed frame can be displayed on any display device of given resolution. The other algorithms such as image segmentation and edge detection are discussed in the research work.

1.3 INTRODUCTION TO VIDEO/IMAGE PROCESSING

Digital Images are produced by a variety of physical devices, including still and video cameras, X-ray devices, electron microscopes, radar, and ultrasound (Bill Silver 2000). The application of image processing includes a variety of fields and in almost every area of human activities. Digital images are used for a variety of purposes, including entertainment, medical, business (e.g. documents), industrial, military, civil (e.g. traffic), security, and scientific (Geoff Dougherty). Digital image processing is applied in the fields of computer vision, face detection, feature detection, lane departure warning system, non-photorealistic rendering, medical image processing, microscope image processing, morphological image processing, remote sensing, telecommunication, etc (Byeong-Hokang 2007). Video/Image processing on mobile device is the application of telecommunication which involves pre-processing and post processing of images. Low level and mid level image processing algorithms can be built easily on mobile device to make the device user friendly which in turn increases the ARPU.

With all this computing power available, it is required to integrate more functionality in the classical cellular telephone. A single unit will have the computing power to be a telephone, a database, a general purpose computer, a gaming machine, a music player, or a digital camera. Therefore, the future of mobile devices is to move from audio-centric to data-centric architectures. It has been projected that until 2010, up to 30,000 million devices will be connected to the global data network [Courtesy: Clark 2002] and most of them will be embedded systems.
System designers need to accomplish real time video and image processing in a small device that consumes minimal power. High quality video and imaging performance requires sophisticated built-in algorithms. Solutions developed for imaging markets thus tend to be balanced toward one of the parameters more than the other (either moderate cost, moderate performance systems with lower power requirements or high cost, high performance systems with increased power requirements). Hence it is required to use the optimized algorithms for the mobile devices to go for moderate cost.

The processor targeted for mobile-handset-camera applications involves the algorithms such as Video/Image compression, Video/Image segmentation, Video/Image scaling, Video/Image enhancement, Video/Image restoration, etc. High quality image processing algorithms generally requires multistage complex algorithms. Complexity and data throughput are the biggest reasons for adopting a hardware approach with Digital Signal Processing (DSP) based architecture or a Field Programmable Gate Array (FPGA) based architecture designs for mobile applications. The mobile handset is one of the leading capture platforms for video content. The mobile handsets are capable of storing, decoding longer video clips with improved resolution. The main limitations associated with the mobile platforms include the limitations in terms of performance, power and memory resources. Also the limitations faced by all mobile application programmers are limited speed control processing units, limited bandwidth and limited memory. Because of the constrained resources of power and cost, and the real-time computation requirements, the processors for use in mobile applications posses a number of distinct characteristics. Examples of such characteristics are limited programmability, high I/O (Input Output) to computation ratio and stream data processing.

1.4 HARDWARE IMPLEMENTATION OF ALGORITHMS

DSP are Microprocessors designed to perform Digital Signal Processing (Mathematical manipulations) of digitally represented signals. Most DSP share some
common basic features to support high performance, repetitive, numerically intensive tasks. As communications get digital there is an increased use of image and video processing in embedded multimedia applications. DSP provide excellent computing platforms for these applications not only due to their superior signal processing performance compared to other processor architectures, but also because of the high levels of integration and very low-power consumption (Texas Instruments 2006) (Klaus Illgner 2000).

![Figure 1.1: Relative power consumption, clock frequency and silicon area of different video coding hardware architectures in H.264 decoding (VGA@30fps)](image)

The three most common solutions used to implement video coding on mobile chips are hardwired (HW) video codecs, video-optimized DSPs, and general-purpose Reduced Instruction Set Computing (RISC) processor cores (often enhanced with SIMD hardware). The level of parallelism is highest in the HW codecs, and lowest in the RISC-based software codecs. Programmable architectures (DSP and RISC) mainly differ in the amount of functional unit level parallelism they offer, and both lack module level parallelism, unless an awkward multi-core solution is used. The figure 1.1 shows the relative power consumption, clock frequency and silicon area of different video coding hardware architectures in H.264 decoding.
The selection of the right DSP for the specific applications depends on
arithmetic format (fixed or floating point), data width, speed, memory organization,
ease of development, multiprocessor support, cost, power consumption &
management. TI’s Open Multimedia Application Platform (OMAP) is a category of
proprietary microprocessors that has capabilities for portable and mobile multimedia
applications and that is developed by TI. Many mobile phones use OMAP
microprocessors, including most of Nokia's N-series range. Phones known to use
OMAP include the N90, N91, N92, N95, N82, E61, E62, E63, E90, N900, Motorola
Droid X, and many other Nokia and Samsung devices. Some of the processors in the
OMAP family contain a dual-core architecture consisting of both a general-purpose
host Advanced RISC (Reduced Instruction Set Computing) Machines (ARM)
processor and one or more DSP. The DSP featured is commonly a variant of the TI’s
TMS320 series.

The OMAP family consists of three product groups classified by performance
and intended applications such as high performance, basic multimedia & modem and
applications. DM642 Evaluation Module (EVM) is a low-cost high performance
Video & Imaging development platform designed for application development and
evaluation of multi-channel, multi-format digital and other future proof applications.
TMS320DM642 is a fixed point processor based on Very Long Instruction Word
(VLIW) architecture and it is a good choice for Video/Image Processing applications.

To be able to bring high-definition video to the hands of users, mobile device
designers must optimize the power consumption of all components. While the display
and wireless transmitter consume a large chunk of the total power, every component
of a mobile system must be evaluated for its contribution to the power budget.
Application and baseband processors must meet stringent power requirements to win a
socket on a mobile handset's circuit board. Video coding is usually the most power-
hungry application a mobile processor has to run. Therefore, minimizing the
maximum power consumption figure of a chip often concentrates on finding the most
power-efficient way to implement video processing algorithms.
Digital mobile chips are generally implemented in low-voltage, low-leakage CMOS processes to minimize the chip's power consumption both in active and standby modes. The drawback of low-power process technologies is its slow operation. Therefore the processing architecture must execute the algorithms in a small number of clock cycles, as high clock-frequencies cannot be used due to long gate delays. This can be achieved by designing an architecture that provides both module and functional unit level parallelism. In other words, the processing architecture has to execute different parts of an algorithm (e.g. IDCT, motion compensation, de-blocking filter, etc in an MPEG 4 video decoder), and different operations within each module (additions, multiplications, memory accesses etc) concurrently in a pipelined manner. Using an extremely parallel architecture makes it possible to use a low clock frequency which in turn makes it possible to use a slow, low-leakage silicon process.

1.5 VIDEO/IMAGE PROCESSING ALGORITHMS

Digital video compression techniques play a key role in recent multimedia communications. The limitation of bandwidth in communication channels and storage media demands more efficient video coding methods. On the other hand, introducing new applications and advances in multimedia technology demands video coding methods to include more complex and advanced features. Therefore, standardization of the video compression techniques is essential. Video compression (Video coding) is the process of compacting or condensing a digital video sequence into a smaller number of bits. ‘Raw’ or uncompressed digital video typically requires a large bit rate and compression is necessary for practical storage and transmission of digital video.

The computational algorithms which can be provided on a mobile phone include Video/Image compression, Video/Image segmentation, Video/Image scaling, edge detection, etc. The different computational algorithms considered provides non-
core services in addition to standard voice calls which is benefit to both the subscriber and service provider.

1.5.1 VIDEO DECOMPRESSION USING H.264 DECODER

Broadcast television and home entertainment have been revolutionized by the advent of digital TV (Television) and Digital Versatile Disc (DVD) video. These applications are made possible by standardizing the video compression technology. For better compression performance, advanced standards such as MPEG-4 and H.263 are introduced. To provide better compression of video compared to previous standards, H.264/MPEG-4 part 10 video coding standards was developed by the Joint Video Team (JVT) consisting of experts from Video Coding Experts Group (VCEG) and Moving Pictures Experts Group (MPEG). H.264 fulfills significant coding efficiency, simple syntax specifications, and seamless integration of video coding into all current protocols and multiplex architectures. Thus H.264 can support various applications like video broadcasting, video streaming, and video conferencing over fixed and wireless networks and over different transport protocols. The coding efficiency advantages of H.264, however, come at the expense of higher computational complexity.

Emerging video coding standards like H.264 achieves significant advances in improving video quality, reducing bandwidth, but at the cost of greatly increased computational complexity at both the encoder and the decoder (Gary.J.Sullivan 2005). Playing encoded videos produced by such compression standards are becoming more popular in mobile applications. H.264 achieves significant advances in improving video quality reducing bandwidth, but at the cost of greatly increased computational complexity at both the encoder and the decoder. Therefore, recently there is growing interest in complexity (power) aware video coding solutions. The High Definition DVD standard (HD-DVD) contains H.264 as one of the three mandatory video formats, next to MPEG-2 and Windows Media Video 9 (WMV9), which is a H.264
variant. Another very close variant, Sorenson Video 3 (SVQ3) is the standard for the
distribution of movie trailers over the Internet. The Digital Video Broadcasting (DVB)
family of digital video infrastructure standards is also adopting H.264 for usage in the
new DVB-Handheld (DVB-H) sub-standard which provides low resolution, low
bandwidth video broadcasting for mobile devices.

The main objective of the emerging H.264 standard is to provide a means to
achieve substantially higher video quality compared to the other existing video coding
standards (Mustafa Parlak 2008). Profiling is done to estimate the computational
complexity and memory requirements. Profiling is to determine the running time of a
block of code.

1.5.2 VIDEO SEGMENTATION ALGORITHM

The progress of multimedia computing and applications created a remarkable
growth in the number of digital images, both still images and video sequences in
computer systems and networks. The need for efficient accessing and processing this
vast amount of pictorial information is evident. One of the most important issues in
this field is computationally efficient representation and comparison of image content.
Segmenting multimedia data streams is a fundamental problem with many
applications. Properly segmented streams can be better organized and reused.
Segmentation algorithms can serve the important function of helping summarize this
mass of material. Video segmentation is required to support object-oriented video
compression technology, like MPEG Layer-4 (MPEG-4) standard (Core, Main
Profile) and to support content-based video processing application (Irena Koprinska
2001). The main intention is to focus on a video, which may be produced video or raw
video. Examples of produced video are news, movies and training videos. Examples
of raw video are video of meetings, surveillance records, and wearable personal video
cameras.
Video segmentation refers to partitioning video into spatial, temporal, or spatiotemporal regions that are homogeneous in some feature space. It is an integral part of many video analysis and coding problems, including (i) video summarization, indexing, and retrieval, (ii) advanced video coding, (iii) video authoring and editing, (iv) improved motion (optical flow) estimation, (v) 3D (3-Dimension) motion and structure estimation with multiple moving objects, and (vi) video surveillance/understanding (Elisa Drelie 2006). As with any segmentation problem, effective video segmentation requires proper feature selection and an appropriate distance measure. Different features and homogeneity criteria generally lead to different segmentations of the same video, for example, color, texture, or motion segmentation. Furthermore, there is no guarantee that any of the resulting automatic segmentations will be semantically meaningful, since a semantically meaningful region may have multiple colors, multiple textures, and/or multiple motions. Although semantic objects can be computed automatically in some well-constrained settings, for example, in video surveillance systems, where objects can be extracted by change detection and background subtraction when the camera is stationary, in general, semantic object segmentation requires specialized capture methods (chroma-keying) or user interaction.

Specific video segmentation methods should be considered in the context of the requirements of the application in which they are used. Factors that affect the choice of a specific segmentation method include the following:

1. Real-time performance
2. Precision of segmentation
3. Scene complexity

Video segmentation algorithm can be classified in to

1. Edge information based algorithm
2. Image segmentation algorithm
3. Change detection algorithm (Yu-When Huang 2002)
There are several algorithms available for segmentation of the video. The genetic segmentation algorithm is advantageous over the other methods for segmentation such as clustering, bilayer segmentation of video, wavelet based methods, eigen value based methods, etc. The genetic mechanism is independent of the prescribed evaluation function and can be tailored to support a variety of characterizations based on heuristics depending on genre, domain, type, etc. The evolutionary algorithms are naturally suited for doing incremental segmentation that can be applied to streaming media. It can support dynamically updated segmentation that adapts to usage patterns, like adaptively increasing the likelihood that frequently accessed points will appear as segment boundaries.

1.5.3 IMAGE COMPRESSION ALGORITHM

Image compression is minimizing the size in bytes of a graphics file without degrading the quality of the image to an unacceptable level. The reduction in file size allows more images to be stored in a given amount of disk or memory space. It also reduces the time required for images to be sent over the internet or downloaded from web pages. Compressing an image is significantly different than compressing raw binary data. Commonly available software can be used to compress images, but the result is less than optimal. This is because images have certain statistical properties which can be exploited by encoders specifically designed for them. Transmission time can be estimated for the given MODEM. The basic speed available in modern mobile handsets (3G, HSPA) or even in GPRS (2G system) is 56 kbps. For 3G the speed up to 200 kbps is available.

Also, some of the finer details in the image should be sacrificed for the sake of saving bandwidth or storage space. The table 1.1 shows the qualitative transition from simple text to full-motion video data and the disk space, transmission bandwidth, and transmission time needed to store and transmit such uncompressed data. The examples
above clearly illustrate the need for sufficient storage space, large transmission bandwidth, and long transmission time for image, audio, and video data. At the present state of technology, the only solution is to compress multimedia data before its storage and transmission, and decompress it at the receiver for play back. For example, with a compression ratio of 32:1, the space, bandwidth, and transmission time requirements can be reduced by a factor of 32, with acceptable quality (A.Messaoudi 2005).

Table 1.1 Multimedia data types and uncompressed storage space, transmission bandwidth, and transmission time required.

<table>
<thead>
<tr>
<th>Multimedia Data</th>
<th>Size/Duration</th>
<th>Bits/Pixel or Bits/Sample</th>
<th>Uncompressed Size (B for bytes)</th>
<th>Transmission Bandwidth (b for bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A page of text</td>
<td>11” x 8.5”</td>
<td>Varying resolution</td>
<td>4-8 KB</td>
<td>32-64 Kb/page</td>
</tr>
<tr>
<td>Telephone quality speech</td>
<td>10 sec</td>
<td>8 bps</td>
<td>80 KB</td>
<td>64 Kb/sec</td>
</tr>
<tr>
<td>Grayscale Image</td>
<td>512 x 512</td>
<td>8 bpp</td>
<td>262 KB</td>
<td>2.1 Mb/image</td>
</tr>
<tr>
<td>Color Image</td>
<td>512 x 512</td>
<td>24 bpp</td>
<td>786 KB</td>
<td>6.29 Mb/image</td>
</tr>
<tr>
<td>Medical Image</td>
<td>2048 x 1680</td>
<td>12 bpp</td>
<td>5.16 MB</td>
<td>41.3 Mb/image</td>
</tr>
<tr>
<td>SHD Image</td>
<td>2048 x 2048</td>
<td>24 bpp</td>
<td>12.58 MB</td>
<td>100 Mb/image</td>
</tr>
<tr>
<td>Full-motion Video</td>
<td>640 x 480, 1 min (30 frames/sec)</td>
<td>24 bpp</td>
<td>1.66 GB</td>
<td>221 Mb/sec</td>
</tr>
</tbody>
</table>

A common characteristic of most images is that the neighboring pixels are correlated and therefore contain redundant information. The foremost task then is to find less correlated representation of the image. Two fundamental components of compression are redundancy and irrelevancy reduction. Redundancy reduction aims at removing duplication from the signal source (Video/Image). Irrelevancy reduction omits parts of the signal that will not be noticed by the signal receiver, namely the Human Visual System (HVS). In general, three types of redundancy can be identified:
- Spatial Redundancy or correlation between neighboring pixel values.
- Spectral Redundancy or correlation between different color planes or spectral bands.
- Temporal Redundancy or correlation between adjacent frames in a sequence of images (in video applications).

Image compression research aims at reducing the number of bits needed to represent an image by removing the spatial and spectral redundancies as much as possible. Since the focus is only on still image compression, temporal redundancy is not considered. Image compression can be classified into lossless and lossy compression or predictive and transform coding techniques. The transform coding methods give a better compression ratio than other methods and hence they are better suited for mobile applications. The transform coding methods involve greater computations and hence it is required to reduce the computation complexity. Transform coding is used to convert spatial image pixel values to transform coefficient values. Image can be compressed by transforming its pixel to a representation where they are decorrelated. Lossy compression can be achieved by quantizing the transformed values. The decoder inputs the transformed values from the compressed stream and reconstructs the original data by applying the inverse transform. The term decorrelated means that the transformed values are independent of one another. If the transformation of the image is done to a representation where the pixels are decorrelated, redundancy of the image is eliminated and the image has been fully compressed.

The different transform coding techniques used for image compression includes Discrete Cosine Transform (DCT), Haar transform, Singular Value Decomposition (SVD), Slant transform, Hadamard transform, Kahunen Loeve Transform (KLT), etc (Dr.Edel Garcia 2006), (Andrew.B.Watson 1994). Among the different transform coding methods listed discrete cosine transform, wavelet transforms and dimensionality reduction techniques are investigated in this proposed research work (Sindhu.M 2009), (Shivali.D.Kulkarni 2008), (Mr.T.Sreenivasulu
Reddy 2007), (Sathish.K.Singh 2010). The suitability of the transform is due to energy compaction property (kamrul Hasn Talukdar 2007). Also, suitability of the transforms is due to subjective quality of the decompressed images in terms of PSNR (Peak signal to Noise Ratio) and quality index, computation time and energy compaction property.

1.5.4 IMAGE SCALING ALGORITHM

One of the important image processing applications on a cell phone is image scaling. Image scaling is generally required to fit the image on the display devices. Image scaling is the process of changing the size of a digital image. Scaling is a non trivial process that involves a tradeoff between speed, smoothness and sharpness. Image scaling can involve either sub sampling (reducing or shrinking an image) or zooming (enlarging an image).

Adjustment of the scale in an image is often desirable either to zoom in on details by magnification or to zoom out to gain an overall overview by reduction. The figure 1.2 indicates the example of image scaling. The equation for interpolation is
\[ f(x, y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} g(i, j) h(x - i, y - j) \]  

(1.1)

Where \( g(x, y) \) is the output image, \( f(x, y) \) is the input image and \( h(x, y) \) represents the interpolation function.

For the case of scaling up (image enlargement), the methods used involve first resizing the image by padding with zeros, followed by convolution with a two-dimensional, spatially invariant linear filter as given by equation 1.1. The filter performs a smoothing function, which dampens out high frequencies in the reconstructed image. Since the image can be represented using matrix which represents discrete pixel values, filter can be used in the discrete domain. Hence each filter can be represented by the impulse function \( h \), which is also called the mask of the filter. Smoothing filter is applied in order to minimize aliasing problems in the target image. Simple integer scaling is accomplished by expansion or reduction of the horizontal and vertical dimensions separately with respect scale factor \( k_x \) and \( k_y \). Magnification requires a replication each existing pixel \( k_x \) times and \( k_y \) times to fill in additional columns and rows of the magnified image. Image scaling can be performed with image interpolation or without image interpolation algorithms (Isahan Pendharkar 2005), (Shuai Yuan 2005), (Ran Gao 2009). The figure 1.3 depicts the image scaling with and without interpolation algorithms. Interpolation is the process of determining the values of a function at positions lying between its samples. It can be achieved by fitting a continuous function through the discrete input samples. The image quality highly depends on the used interpolation technique.

![Figure 1.3: Image scaling with and without interpolation](image1.png)
The interpolation techniques are divided into two categories, deterministic and statistical interpolation techniques. The difference is that deterministic interpolation techniques assume certain variability between the sample points, such as linearity in case of linear interpolation. Statistical interpolation methods approximate the signal by minimizing the estimation error. This approximation process may result in original sample values not being replicated. The speed and quality of transformation of the new pixel values from old pixel values due to scaling depends on the type of interpolation algorithms (Tinku Acharya 2007), (Emil Domic 2007).

The image scaling with interpolation can be achieved by using different interpolation algorithms. The algorithms can be used for image scaling with interpolation includes nearest neighbor interpolation, bilinear interpolation, bicubic interpolation, display/aperture interpolation, Lagrange interpolation, Spline interpolation etc. Among the interpolation algorithm listed the more suitable algorithm for image scaling are nearest neighbor and bilinear interpolation algorithms. The suitability of the algorithm is due to the complexity involved in computation of the new pixel values and memory required to fit the algorithm on the mobile devices.

Interpolation methods differ in their mathematical description of a “good” interpolated image. Although it is difficult to compare methods and judge their output, basic criteria for a good interpolation methods are discussed. The first 8 are visual properties of the interpolated image and the last is a computational property of the interpolation method.

- **Geometric Invariance:** The interpolation method should preserve the geometry and relative sizes of objects in an image. That is, the subject matter should not change under interpolation.
- **Contrast Invariance:** The method should preserve the luminance values of objects in an image and the overall contrast of the image.
- **Noise:** The method should not add noise or other artifacts to the image, such as ringing artifacts near the boundaries.
- **Edge Preservation**: The method should preserve edges and boundaries, sharpening them where possible.
- **Aliasing**: The method should not produce jagged or “staircase” edges.
- **Texture Preservation**: The method should not blur or smooth textured regions.
- **Over smoothing**: The method should not produce undesirable piecewise constant or blocky regions.
- **Application Awareness**: The method should produce results appropriate to the type of image and order of resolution. For example, the interpolated results should appear realistic for photographic images, but for medical images the results should have crisp edges and high contrast. If the interpolation is for general images, the method should be independent of the type of image.
- **Sensitivity to Parameters**: The method should not be too sensitive to internal parameters that may vary from image to image.

## 1.6 SURVEY OF LITERATURE

Cisco is predicting in its Cisco Visual Networking Index (VNI) Forecast: 2008–2013 report that the advent of 4G services will lead to an explosion of new video applications that will transform telecom service providers into what the company is dubbing “experience providers.” Cisco also predicts that high-speed wireless broadband technologies such as WiMAX and Long Term Evolution (LTE) will make high-definition video streaming widely available on both fixed and mobile devices. The upshot of this is that telecom providers will move into the cable companies’ traditional territory by offering more comprehensive video services. Cisco’s recent study on Internet traffic trends projects that by 2013: 64% of mobile data traffic will be for video, 19% for data services, 10% for peer-to-peer, and 7% for audio. The study also says that the projected video traffic will increase four-fold between now and 2012. This coming increase in available video services and applications mirrors what has happened every time new Web services expand their bandwidth capabilities (White paper 2009).
Research work discusses applications of image processing as image scaling, image compression, image enhancement, edge detection, image segmentation. Research work also extended to applications such as video compression, video segmentation. Solutions for mobile devices require a fundamentally different approach than traditional vision techniques that run on traditional computers, so we consider user-device interaction and the fact that these algorithms must execute in a resource constrained environment. The thesis makes contributions related to efficient implementation of Video/Image processing techniques. A detailed literature survey has been conducted in the domain of advanced video coding implementation on DSP, segmentation and summarization of video, image compression and image scaling.

The implementation of the H.264 decoder on different processors to estimate the complexity involved in implementation of the decoder. A literature survey on implementation of H.264 decoder on DSP is conducted and the outcome of the few survey papers is listed. (Sourabh Rungta 2009) proposes an algorithm to develop software which implement H.263 decoder and encoders so that the software implementation must be highly optimized to achieve “reasonable” video quality. (Seong Mo Park 2006) discusses the design of a single chip video decoder called advanced mobile video ASIC (A-MoVa) for mobile multimedia applications. (Tsu-Ming Liu 2005) presents a low power H.264/AVC video decoder LSI for mobile applications. (Yasir Mahe 2005) discusses the three implementations of an H.264/AVC baseline decoder using different leading high-end Digitals Signal Processors. H.264 baseline real-time video decoder implementation based on TMS320DM642 and JM86 test code is described in (Liu, Hui 2008). This chip uses mixed hardware/software architecture to improve both its performance and its flexibility. The implementation of a Baseline Profile and main profile H.264 decoder based on a DM642 digital signal processor is discussed (F.Pescador 2007). Study and analyze the computational complexity of a software-based H.264/AVC baseline profile decoder is presented in the paper (Michael Horowitz 2003). Latest generation DSPs are becoming more efficient, being able to improve their forerunners while
reducing their internal memory size to lower the cost. An H.264 video decoder based on a latest generation DSP is described (F. Pescadores 2009). (Shing Yu Shih) proposes efficient hardware architecture for the deblocking filter function in H.264/AVC. (Gulistan Raja) proposes in-loop deblocking filter and test results show that H.264’s in-loop deblocking filter can significantly reduce the blocking artifacts. The H.264/AVC standard defines an optional in-loop deblocking filter. The effect of this filter on subjective video quality is investigated. Two efficient and low power H.264 deblocking filter (DBF) hardware implementations that can be used as part of an H.264 video encoder or decoder for portable applications is discussed (Mustafa Parlak 2008). (Zhenhua Xu 2009) presents an implementation of Exponential Golomb (Exp-Golomb) code which is one of the key techniques in AVS video compression standard on Storm stream processor. Experimental results show that the real-time requirement of high-resolution video application can be achieved.

For some applications, it is required to segment the video and summarize the frames to select only few frames in the video clip. The different algorithms for segmentation of the video are studied. Temporal video segmentation for real-time key frame extraction is discussed (Calic, J 1993). A temporal video segmentation method proposed is based on the detection of shot abrupt transition and gradual transition, and then takes into account the conditions of user terminals, which could generate different video summarization for each user (Chen Yinzi 2010). Implementation of image segmentation using GA involves defining a fitness evaluation function, designing a ‘Population’ (set of chromosomes), defining genetically inspired operators such as crossover and mutation to evolve new population and deciding the termination of the evolutionary search for the optimal solution (Aravind.I 2002). The genetic segmentation algorithm for video is defined and the evolutionary nature of genetic algorithms offers a advantage by enabling incremental segmentation (Patric Chiu 2000). In the computational process, the improved GA adjusts crossover probability and mutation probability automatically according to the variance between the target and background, thus overcoming the problems of Simple Genetic Algorithm (Lei

The schemes for using learning in video analysis tasks like content based filtering and shot summarization is discussed (Koustav Bhattacharya 2004). A new approach for automatic video scene segmentation and content based indexing approach detects video shots and builds a collection of key frames and representative frames. Scene segmentation and video indexing are based on a temporally windowed principal component analysis of a sub-sampled version of the video sequence (Keesook J. Han 1997). A new method is presented for extracting key frames. Firstly, an improved histogram matching method is used for video segmentation. Secondly, the key frames are extracted utilizing the features of I-frame, P-frame and B-frames (Guozhu Liu 2009). (Paulo Lobato Correia 2003) discusses the problem of video segmentation quality evaluation, proposing evaluation methodologies and objective segmentation quality metrics for individual objects as well as for complete segmentation partitions. (Tsong-Yi Chen 2009) proposes an efficient real-time video object segmentation algorithm based on change detection and background updating. A video segmentation method using a histogram-based fuzzy c-means (HBFCM) clustering algorithm is discussed (Chi-Chun Lo 2001). The concept of a priority curve associated with a video and it can be used to create a summary (of a desired length) of any video (M. Fayzullin 2004). (Chi-Chun Lo 2003) propose a histogram-based moment-preserving (HBMP) clustering algorithm for segmenting video data. This algorithm is a hybrid of the shot change detection approach and the clustering approach (Yasira Beevi.C.P 2009) presents a video segmentation algorithm for MPEG-4 camera system with change detection, background registration techniques and real time adaptive threshold techniques

A variation to the SVD based image compression technique is proposed to compress the decoded frames. The variation can be viewed as a preprocessing step in which the input image is permuted as per a fixed, data independent permutation, after which it is fed to the standard SVD algorithm (Abhiram Ranade 2007). The DCT is
used to transform the highly correlated blocks of the YCbCr components, while the SVD is used to transform the low correlated blocks (Y. Wongsawat, 2004). A good compromise between the quality and the compression rate factors that are achieved when processing images by the DCT technique are discussed (A. Messaoudi 2005).

The decompressed image is displayed on a mobile phone of any resolution using scaling algorithms. Common approach is to start with an original image, generate a lower resolution version of original image by downscaling, and then use different interpolation methods to magnify low resolution image. After that original and magnified images are compared to evaluate difference between them using different picture quality measures. For fair comparison all these parameters need to be considered (Emil Dumic 2007) (Ryuji Matsuoka 2008). A log-polar neighbor model is adopted, utilizing the feature of applying larger weights to pixels at the center of the interpolation region and logarithmically decreasing weights to pixels away from the center (A. Amanatiadis 2007). A comparison of different interpolation techniques both in the spatial and frequency domains. Various methods are presented, from a very simple linear interpolation to more complex cubic convolution (Miklos Poth 2010). A parallel between different resampling algorithms is made, and as basis for this comparison are used three methods: Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR) and Subjective Evaluation (SE) (Slavy Georgiev Mihov, 2005).

A fast algorithm for image interpolation is proposed for real-time enlargement of video images is discussed (Mei-Juan Chen 2005). A novel method of nonlinear scaling is proposed to overcome the compatibility problem (Shi Zaifeng 2008). An adaptive resampling algorithm for zooming up images is based on analyzing the local structure of the image and applying a near optimal and least time-consuming resampling function that will preserve edge locations and their contrast (A.M. Darwish 1997). A method is proposed to take into account information about discontinuities or sharp luminance variations while doubling the input picture (S. Battiato 2002). Three popular methods such as bilinear interpolation, cubic convolution and cubic convolution are investigated (Ryuji Matsuoka 2008) (Kian Kee
The different measures can be used to compare the different algorithms of interpolation (Zhou Wang 2002), (Satu Jumisko 2005), Niranjan Damera Venkata (2000), (Ismail Avcibas 2002).

1.7 PROPOSED SOLUTION

The research work carried out is to investigate the complexity involved in the video processing services. The advanced video coding standard H.264 decoder is implemented on a hardware platform such as TI DSP (TMS320C642 DSP clocked at 600 MHz) to measure the computation complexity and memory requirements. The DM642 processor is the low-cost high performance video & imaging development platform designed for application development and evaluation of multi-channel, multi-format digital applications. To make use of H.264 baseline decoder on mobile applications, the resources are limited in terms of memory and computational requirements. Hence it is required to compress the data and also analyze the data based on summarization after segmentation. All the decoded frames may not be necessary and only selected frames are needed to view the highlights of an event. Thus the research work is extended to video segmentation and a method is proposed to modify the genetic algorithm for video segmentation. The H.264 standard removes the temporal redundancy. The spatial redundancy can be achieved by using image compression. Image compression is required to reduce the memory space required to store the image. Lossy image compression gives higher compression ratio without significantly affecting the viewer’s perception of visual quality compared to lossless image compression. Different transform coding methods such as discrete cosine transform (DCT), discrete wavelet transform (DWT), KL transform, and SVD etc can be used to achieve lossy image compression. SVD is an algorithm with higher level of complexity to implement but it has very good energy compaction property. The complexity in computation is reduced using a novel approach. Research work is carried out to preprocess the SVD and is named in this research work as modified SVD. The decompressed image can be of any resolution and to display the image on a
mobile device, image scaling is very essential. The image scaling can be achieved using interpolation techniques and the research work discusses a novel method considering computation complexity and the quality in terms of PSNR and quality index for image scaling.

1.8 ORGANIZATION OF THE DISSERTATION

The thesis is organized into as follows.

The chapter 1 contains introduction, literature survey and motivation for the research. The chapter briefs about the background for the research, problem formulation, algorithm of Video/Image processing on mobile devices and problem solution. The background for the research explains about the introduction to Video/Image processing, applications of Video/Image processing in particular application in communication fields. Also the chapter gives insight into Video/Image processing on mobile devices, the hardware implementation of the different algorithms, hardware details in particular DSP evaluation module DM 642.

The Chapter 2 discusses the advanced video coding standard H.264 decoder used for decompressing the H.264 encoded video frames. Porting of H.264 decoder on TIDSP is done to measure the computational complexity and memory requirements. The profiling is done to estimate the computational complexity and to estimate the memory requirements for each block of H.264 decoder.

The Chapter 3 presents the video segmentation and summarization aspects. Segmentation and summarization is done on the decoded frames to view only key frames. The chapter discusses about the genetic algorithm for image segmentation and summarization of the frames decoded. The method is proposed to modify the genetic algorithm and the proposed method is compared with the genetic algorithm. Also, the
algorithm is compared with original genetic algorithm for different set of input images.

The Chapter 4 presents the image compression using transform coding techniques. The chapter discusses the singular value decomposition method for compressing the images which is compared with DCT and Haar wavelet transforms. A method is proposed to preprocess the SVD to reduce computation complexity involved in computing the SVD. The comparison of the method proposed is done with SVD and comparison is done using different parameters. The suitability of the method is highlighted for the given mobile applications.

The Chapter 5 presents the image scaling using interpolation techniques to fit the image on a display device. The method is proposed to scale the image and the proposed method is compared with the existing interpolation techniques such as nearest neighbor and bilinear. Comparison is done by considering different parameters and the suitability of the proposed method is highlighted.

Finally, conclusions and directions for future research are listed in chapter 6.

In the appendix A, additional details about the hardware implementation are discussed. The introduction to mobile device, Video/Image processing on mobile device and hardware details about the TI processor TMS320CDM 642 are discussed in this section.

In the appendix B, a detailed investigation of image segmentation algorithms is conducted by porting the algorithm on DSP. In particular image segmentation using edge detection method and segmentation algorithms are discussed. The segmentation algorithms are discussed in detail using different edge detection operators. The comparison of the different edge detection operators is conducted using different parameters.

In the Appendix C, answer to the queries raised by the examiners (both Indian and foreign examiners) are discussed.