IMPLEMENTATION OF THREE STEP SEARCH (TSS) ALGORITHM IN MOTION ESTIMATION METHOD USING MATLAB

NAZIHAN BINTI AZIZIUDDIN

This report is submitted in partial fulfillment of the requirement for the award of Electronic Engineering (Computer Engineering) With Honours

Faculty of Electronic and Computer Engineering
Universiti Teknikal Malaysia Melaka

May 2008

© Universiti Teknikal Malaysia Melaka

UNIVERSTI TEKNIKAL MALAYSIA MELAKA
FAKULTI KEJURUTERAAN ELEKTRONIK DAN KEJURUTERAAN KOMPUTER

BORANG PENGESAHAN STATUS LAPORAN
PROJEK SARJANA MUDA II

Tajuk Projek : IMPLEMENTATION OF THREE STEP SEARCH (TSS) ALGORITHM IN MOTION ESTIMATION METHOD USING MATLAB

Sesi Pengajian : 2007/2008

Saya NAZIHAN BINTI AZIZIUDDIN (HURUF BESAR)

mengaku membenarkan Laporan Projek Sarjana Muda ini disimpan di Perpustakaan dengan syarat-syarat kegunaan seperti berikut:

1. Laporan adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan laporan ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (√) :

☐ **SULIT*** (Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)

☐ **TERHAD*** (Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

☒ **TIDAK TERHAD**

Disahkan oleh:

(TANDATANGAN PENULIS)

Alamat Tetap : KM 24, Kampung Air Lanjut,

78 000 Alor Gajah, Melaka

(COP DAN TANDATANGAN PENYELIA)

REDZUAN B ABDUL MANAP
Ketua Jabatan (Kej Telekomunikasi)
Fakulti Kej Elektronik dan Kej Komputer (FKEKK)
Universiti Teknikal Malaysia Melaka (UTeM)
Karung Berkunci 1200,
Ayer Keroh, 75450 Melaka

Tarikh: 9.5.2008

Tarikh: 9/5/2008

© Universiti Teknikal Malaysia Melaka

"I hereby declare that this report is result of my own work except for quotes as cited in the references."

Signature : ....

Name : NAZIHAN BINTI AZIZIUDDIN

Date : 9 MAY 2008

© Universiti Teknikal Malaysia Melaka

"I hereby declare that I have read this report and in my opinion this report is sufficient in terms of the scope and quality for the award of Bachelor of Electronic Engineering (Computer Engineering) With Honours."

Signature : ....

Supervisor's Name : EN. REDZUAN BIN ABDUL MANAP

Date : 9 MAY 2008

© Universiti Teknikal Malaysia Melaka

To my beloved mother, brothers and sisters

© Universiti Teknikal Malaysia Melaka

ACKNOWLEDGMENT

First and foremost, I would like to express my heartily gratitude to my supervisor, En Redzuan bin Abdul Manap for the guidance and enthusiasm given throughout the progress of this project.

My special dedication to my beloved mother, brothers and sisters who has been so tolerant and supportive all these years. Big thanks for their encouragement and love they have given me and without their help, I would not have reach so far here.

Nevertheless, my great appreciation dedicated to all my friends and colleagues and those whom have involve directly or indirectly with this project. Thank you very much for great support and the encouragement that I received from Nur Adilah, Nur Amalina, Kamaliah Hanim, and Fadlilhisham. Their views and tips are useful indeed.

© Universiti Teknikal Malaysia Melaka

ABSTRACT

The main purpose of this project is to compress video sequence using Block Matching Motion Estimation coding standards. To achieve high compression ratio in video coding techniques many previous fast block-matching algorithms have been proposed and developed. One of these, which are proposed for the implementation in this project, is called Three Step Search (TSS) Algorithm. This implementation is using MATLAB software. While developing this algorithm coding, there is a certain guideline that must be followed in order to ensure the project is completely finished without facing much difficulty. At the end of this project, TSS Algorithm for motion estimation techniques has been successfully implemented. The performance of this algorithm is compared to other common algorithms in terms of Peak Signal to Noise Ratio (PSNR), number of search points required and computational complexity.

© Universiti Teknikal Malaysia Melaka

ABSTRAK

Tujuan utama projek ini adalah untuk memampatkan pengekodan video dengan menggunakan kaedah pengekodan persamaan blok. Untuk mencapai nisbah kemampatan yang tinggi dalam teknik pengekodan video, banyak algoritma blok sepadan telah dicadangkan dan dibangunkan untuk kaedah ini. Salah satu algoritma yang dicadangkan untuk dilaksanakan dalam projek ini adalah Three Step Search (TSS) Algorithm. Pelaksanaan ini menggunakan perisian *MATLAB*. Terdapat garis panduan tertentu yang perlu dipatuhi dalam melaksanakan projek ini bagi memastikan projek ini dapat disiapkan dengan sempurna tanpa berhadapan dengan sebarang masalah. Di penghujung projek ini, algoritma TSS telah Berjaya disimulasikan di dalam MATLAB. Prestasi algoritma ini dibandingkan dengan algoritma yang lain melalui Peak Signal to Noise Ratio(PSNR), bilangan titik pencarian yang diperlukan dan kerumitan dalam pengkomputeran.

© Universiti Teknikal Malaysia Melaka

TABLE OF CONTENT

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

LIST OF TABLE

© Universiti Teknikal Malaysia Melaka

LIST OF FIGURE

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

LIST OF ACRONYMS

BBGDS	Block Based Gradient Descent Search
BMA	Block Matching Algorithm
BMC	Block Matching Compensation
CCB	Cross Center Biased
CCF	Cross Correlation Function
CDS	Cross Diamond Search Algorithm
CS	Cross Search
CSP	Cross Shaped Search Pattern
DCB	Diamond Center Biased
DCT	Discrete Cosine Transform
DS	Diamond Search
FSS	Four Step Search
GOP	Group of Picture
JPEG	Joint Photographic Expert Group
LDSP	Large Diamond Search Pattern
MAE	Mean Absolute Error
MC	Motion Compensation
ME	Motion Estimation
MPEG	Moving Picture experts Group
MSD	Mean Squared Difference
MSE	Mean Squared Error
MV	Motion Vector
NTSS	New Three Step Search

© Universiti Teknikal Malaysia Melaka

OSA	Orthogonal Search Algorithm
OTS	One at a Time Search
PSNR	Peak signal Noise Ratio
SDSP	Small Diamond Shaped Pattern
TDL	Two Dimensional Logarithmic
TSS	Three Step Search
VBSMC	Variable Block Size Motion Compensation
VCD	Video CD

© Universiti Teknikal Malaysia Melaka

CHAPTER I

INTRODUCTION

1.1 Introduction

Nowadays, with the increasing demand of multimedia applications, large efforts are needed for efficient video compression and encoding algorithms. Video compression is use because video sequence consists of much redundant information which must be exploited in order to be stored and transmitted efficiently. To achieve high compression ratio in video coding techniques, a method called Block Matching Motion Estimation is widely used in various coding standards. Motion estimation (ME) plays an important role in motion compensated image sequence coding. This technique is used because it has proven to be an effective technique for exploiting the temporal redundancy in video sequences.

Many previous fast Blocks Matching Algorithms (BMA) have been proposed and developed for this method. For example, the New Three Step Search (NTSS), the Four Steps Search (FSS), the Block Based Gradient Descent Search (BBGDS), the Diamond Search (DS), and Cross Search (CS). These algorithms have different technique and process for video compression.

© Universiti Teknikal Malaysia Melaka

One of these, which is proposed for the implementation in this project, is called Three Step Search (TSS) algorithm. The aim of this algorithm is to find the optimum motion vector (MV) with minimal number of search points along the search process. The performance of this algorithm is compared to other common algorithms in terms of Peak Signal to Noise Ratio (PSNR), number of search points required and computational complexity.

1.2 Objectives

The objectives of this project are:

1) The project aimed to investigate one of the BMA available, known as TSS Algorithm.
2) To implement the algorithm using MATLAB.
3) To analyze its performance compared to other common BMA.

1.3 Problems Statement

There are several techniques used in video coding to achieve high compression. BMA is one of common techniques that is being used. Since there is a lot of BMA fast algorithms have been developed, it is difficult to choose the best algorithm to be implemented.

1.4 Scope of Work

The scopes of work for this project can be divided into two:

© Universiti Teknikal Malaysia Melaka

1) Acquisition of information involving background theory on Motion Estimation, Block Matching Algorithm, MATLAB, Three Step Search (TSS) Algorithm, and video coding techniques.
2) Implementation of Three Step Search (TSS) Algorithm using MATLAB also performance and result analysis.

1.5 Thesis Layout

Chapter 2 consists of theories and background on video motion compensation and ME technique. This is followed by theories of block matching and its algorithms. In this chapter the advantages of using MATLAB are also given.

The overview of the methodology used in this project is described in Chapter 3. Details from the beginning of the project until the end of this project explained in this chapter. Project is started with data acquisition and literature review. Step one in MATLAB is extracting video sequence into frame. This is followed by BMA and implementation of TSS algorithm implementation. Then the predicted frame is constructed followed by performance and results analysis.

Experimental and simulated results are given in Chapter 4. The result is dividing into 2 parts. For the first part, only TSS algorithm is implemented into the standard video sequence. PSNR and search point performance of the algorithm is shown. For the second part the comparison of the TSS algorithm, CDS algorithm and NTSS algorithm are made.

Chapter 5 consists of the discussion of the result obtained and the performance of the algorithm. The performance of each algorithm is compared to determine which is better. This part also has the conclusion of the project and suggestion for future work.

© Universiti Teknikal Malaysia Melaka

CHAPTER II

LITERATURE REVIEW

2.1 Motion Estimation (ME)

ME is a one of the key elements of many video compression schemes. A video sequence consists of a series of frames. To achieve compression, the temporal redundancy between adjacent frames can be exploited. That is, a frame is selected as a reference, and subsequent frames are predicted from the reference using a technique known as ME. The process of video compression using ME is also known as interframe coding [1].

In a sequence of frames, the current frame is predicted from a previous frame known as reference frame. The current frame is divided into macroblocks, typically 16 x 16 pixels in size. This choice of size is a good trade-off between accuracy and computational cost. However, ME techniques may choose different block sizes, and may vary the size of the blocks within a given frame.

Each macroblock is compared to a macroblock in the reference frame using some error measure, and the best matching macroblock is selected. The search is conducted over a predetermined search area. A vector denoting the displacement of the macroblock in the reference frame with respect to the macroblock in the current frame is determined. This vector is known as MV.

© Universiti Teknikal Malaysia Melaka

When a previous frame is used as a reference, the prediction is referred to as forward prediction. If the reference frame is a future frame, then the prediction is referred to as backwards prediction. Backwards prediction is typically used with forward prediction, and this is referred to as bidirectional prediction [2].

2.2 Motion Compensation (MC)

In video compression, MC describes a picture in terms of where each section of that picture came from, in a previous picture. This is often employed in video compression. It can also be used for deinterlacing. A video sequence consists of a number of pictures usually called frames. Subsequent frames are very similar, thus containing a lot of redundancy. Removing this redundancy helps achieve the goal of better compression ratios [1].

A first approach would be to simply subtract a reference frame from a given frame. The difference is then called *residual* and usually contains less energy or information than the original frame. The residual can be encoded at a lower bit-rate with the same quality. The decoder can reconstruct the original frame by adding the reference frame again.

A more sophisticated approach is to approximate the motion of the whole scene and the objects of a video sequence. The motion is described by some parameters that have to be encoded in the bit-stream. The pixels of the predicted frame are approximated by appropriately translated pixels of the reference frame. This gives much better residuals than a simple subtraction. However, the bit-rate occupied by the parameters of the motion model must not become too large.

Usually, the frames are processed in groups. One frame (usually the first) is encoded without MC just as a normal image. This frame is called *I-frame* (intra-coded frame, Motion Picture Expert Group (MPEG) terminology) or *I-picture*. The other frames are called *P-frames* or *P-pictures* and are predicted from the I-frame or P-frame that comes immediately before it. The prediction schemes are, for instance,

© Universiti Teknikal Malaysia Melaka

described as IPPPP, meaning that a group consists of one I-frame followed by four P-frames.

Frames can also be predicted from future frames. The future frames then need to be encoded before the predicted frames and thus, the encoding order does not necessarily match the real frame order. Such frames are usually predicted from two directions, i.e. from the I- or P-frames that immediately precede or follow the predicted frame. These bidirectional predicted frames are called *B-frames*. A coding scheme could, for instance, be IBBPBBPBBPBB [3].

2.3 Variable Block-Size Motion Compensation (VBSMC)

VBSMC is the use of Block Motion Compensation (BMC) with the ability for the encoder to dynamically select the size of the blocks. When coding video, the use of larger blocks can reduce the number of bits needed to represent the MVs, while the use of smaller blocks can result in a smaller amount of prediction residual information to encode. Older designs such as H.261 and MPEG-1 video typically use a fixed block size; while newer ones such as H.263, MPEG-4 Part 2, H.264/MPEG-4 AVC, and VC-1 give the encoder the ability to dynamically choose what block size will be used to represent the motion [4].

2.4 Video Compression

MPEG-1 is an ISO standard developed by the Moving Pictures Expert Group. It was developed for Video CDs (VCD). While VCDs never became popular because movies were broken into several CDs, the standard gained acceptance in a variety of other applications. Most of the MPEG files found on the internet are MPEG-1 files. In addition, the output of digital video cameras is often in an MPEG-1 stream. In short, MPEG-1 is a digital video compression standard which has gained wide acceptance in the past decade [5].

© Universiti Teknikal Malaysia Melaka

Before the information encoded can be applied in an MPEG video, what information is encoded in an MPEG must be thoroughly understood. MPEG seeks to compress digital video as much as possible while still maintaining a high quality picture. The first step in the compression process involves switching from the RGB color space to the YCrCb color space. YCrCb describes a color space where the three components of color are luminance, red chrominance, and blue chrominance. This switch is made because the human eye is less sensitive to chrominance than it is to luminance. Chrominance data can then be sampled at a quarter the rate of luminance data.

The next step in compression involves reducing spatial redundancy. This is done using essentially the same methods as Joint Photographic Experts Group (JPEG). The image is divided into 16 x 16 pixel macroblocks. Each macroblock contains 16 x 16 luminance pixels, and 8 x 8 red/blue chrominance pixels. The luminance block is then split into 4 8 x 8 blocks. Now, there are 6 8 x 8 blocks on which a Discrete Cosine Transform (DCT) is performed. The DCT coefficients are quantized, filtered, and then stored.

The next step in compression is intended to reduce temporal redundancy. The first step in this process is to divide a series of frames into a group of pictures (GOP) and then to classify each frame as either I, P, or B. The usual method is to break a video into GOPs of 15 frames. The first frame is always I frame. In a 15 frame GOP, it is common to have two B frames after the I frame, followed by a P frame, followed by two B frames..

The classification of a frame as I, P, or B determines the manner in which temporal redundancies are encoded. An I frame is encoded "from scratch", just as described above. However, a P frame is encoded by breaking the image into macroblocks, and then using a matched filter, or a similar setup, to match each macroblock to a 16 x 16 pixel region of the last I frame. Once the best match is found, the MV is assigned to that macroblock, and the error between the DCT coefficients of the current macroblock and the region it is being compared to from the I frame is encoded (as DCT coefficients). A B frame differs from a P frame only in that the above step is performed twice, once relating a macroblock in B to a point

© Universiti Teknikal Malaysia Melaka