國立交通大學

資訊工程學系

資訊科學與工程研究所

碩 士 論 文

針對畫質與頻寬限制的串流系統，使用自適性跳畫面機制的初步探討

A Preliminary Study on Adaptive Frame Skipping for Quality-and Rate-Constrained Streaming Systems

研 究 生：林岳進

指導教授：彭文孝 博士

中 華 民 國 九 十 七 年 九 月

針對畫質與頻寬限制的串流系統，使用自適性跳畫面機制的初步探討

A Preliminary Study on Adaptive Frame Skipping for Quality-and Rate-Constrained Streaming Systems

研究生　：林岳進　　Student: Yuei-Chin Lin

指導教授: 彭文孝 博士　　Advisor: Dr. Wen-Hsiao Peng

國 立 交 通 大 學

資 訊 工 程 學 系

資 訊 科 學 與 工 程 研 究 所 碩 士 班

碩 士 論 文

A Thesis
Submitted to Department of Computer Science
College of Electrical and Computer Engineering
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of Master
in
Computer Science

September 2008
Hsinchu, Taiwan, Republic of China

中華民國九十七年九月

國立交通大學

研究所碩士班

論文口試委員會審定書

本校 資訊科學與工程 研究所 林岳進 君

所提論文:

A Preliminary Study on Adaptive Frame Skipping for Quality- and Rate-Constrained Streaming Systems

針對畫質與頻寬限制的串流系統，使用自適性跳畫面機制的初步探討

合於碩士資格水準、業經本委員會評審認可。

口試委員： 林嘉文 彭文孝

[illegible]

指導教授： 彭文孝

所　　長： [illegible]

中 華 民 國 九十七 年 九 月 十二 日

針對畫質與頻寬限制的串流系統，使用自適性跳畫面機制的初步探討

研究生: 林岳進　　　　　　　　　　　　　指導教授: 彭文孝 博士

國立交通大學

資訊工程學系 資訊科學與工程研究所碩士班

摘要

由於人眼對於畫質好壞差異的感受比斷斷續續畫面所帶來的停頓效果較不敏感，大多編碼器都會注重在將可用位元數分配到所有可編畫面上，即使個別畫面會因此得到較少的可用位元而使得畫質變差，也大多不願意將畫面率降低來得到較好的畫質。但是對於監控系統或是一些比較在意畫面品質的攜帶型設備而言，則會對於編碼畫面品質上有一定要求。

因此如何在同時具有畫質與頻寬限制的環境達到最佳化變成我們想要解決的議題。為了解決多重限制的問題，我們首先從了解最佳解的設計出發，再深入研究如何輔以 Lagrange multiplier 的方式來達到位元率-失真交換下(Rate-Distortion Trade-Off)最佳化。

相較於一般流量控制的設計都是只考量頻寬限制的要求，我們的問題勢必要使用動態規劃才能達到最佳位元分配；然而為了降低複雜度，支配線(Dominative Line)與依次精修(Successive Refinement)的方法被提出並且分析其結果。

A Preliminary Study on Adaptive Frame Skipping for Quality-and Rate-Constrained Streaming Systems

Student: Yuei-Chin Lin Advisor: Dr. Wen-Hsiao Peng

Department of Computer Science &
Institute of Computer Science and Engineering
National Chiao Tung University

Abstract

Owing to that human eyes are more sensitive to jerky effect caused by different frame rate rather than quality variance caused by different quality of coded frames, most encoder systems focus on distributing available bit budget among all frames and are not willing to reduce frame rate to obtain better spatial resolution, even the quality of each frame becomes worse for getting less available bits. However, there exists quality requirements to surveillance systems and some mobile devices which care about the quality of each coded frame.

As a result, how to solve the problem with both distortion- and rate-constraint is the main issue in our research. In order to solve this multiple constrained problem, we start by understanding the design of optimal solution; furthermore, we study how to use Lagrangian parameters for rate-distortion trade-off optimization.

Compared with general rate control scheme, which only consider the rate/budget requirement, our problem must use dynamic programming for optimal bit allocation. Nevertheless, in order to reduce complexity, dominative line and successive

refinement methodology are proposed and analyzed.

誌謝

在這兩年的研究生涯中，我首先要感謝指導教授彭文孝老師。從碩一開始進實驗室到碩二畢業的兩年間，老師不僅在研究上給予我專業的指導，也同時教導我研究學問的方法與態度。在此僅向老師至上最高的感謝之意。

這篇論文可以完成，也要感謝交大這個大環境，能夠讓我在這塊領域中討論，並且解決我的疑問。MAPL 就像個大家庭，實驗室提共了豐富的資源與專業的設備，還有良好的研究環境，讓我能夠在這裡順利的完成學業。我也很感謝實驗室的雪婷與敏正、志展同學還有淯紋、鴻志、志鴻學長們，可以共同討論，共同成長。

最後，我也要感謝我的父母林豐獎先生及陳淑女女士與親戚給予的幫助；另外秋盛、育青、孝嶸等摯友們，在我研究艱困的時候給予精神上的支持。在此僅將論文獻給所有幫助過我，陪伴我走過求學生涯的所有師長，朋友、同學和家人。

林岳進

2008 年九月于新竹

A Preliminary Study on Adaptive Frame Skipping for Quality-and Rate-Constrained Streaming Systems

Advisor: Prof. Wen-Hsiao Peng

Student: Yuei-Chin Lin

Institute of Multimedia Engineering

National Chiao-Tung Univeristy

1001 Ta-Hsueh Rd., 30010 HsinChu, Taiwan

August 2008

Contents

Contents i

List of Tables iii

List of Figures iv

1 Research Overview 1

1.1 Introduction 1

1.1.1 Architecture 1

1.1.2 Objective and Formulation 2

1.2 Related Works 3

1.3 Contribution and Organization of Thesis 5

2 Constrained Optimization Problems : Principles and Applications 6

2.1 Background 6

2.2 Integer Programming 7

2.3 Lagrangian Optimization 8

2.3.1 Optimization Problems with Multilple Constraints 9

2.3.2 Optimization Problems with Single Constraint 12

2.4 Complexity Comparison 17

3 Related Works **18**

3.1 KKT Conditions 18

3.2 Viterbi Algorithm with Lagrangian Cost 20

3.3 Variable Frame Rates Encoding 21

3.4 Comparison 23

4 Experiments and Analyses **25**

4.1 Optimal Solution Analyses 25

4.1.1 MSE Weighting Effect 26

4.1.2 Optimal Path Analyses 29

4.2 Proposed Heuristic Solution 32

4.2.1 Coding Frame Selection Methodologies 32

4.2.2 Successive Refinement 34

4.3 Proposed Rate Control Mechanism 36

5 Conclusions and Future Works **39**

Bibliography **42**

List of Tables

2.1 Complexity comparison between DP 17

3.1 Comparison among the proposed problem and related works 23

List of Figures

1.1 Proposed video streaming system architecture. 2
1.2 Proposed adaptive frame coding. 3
2.1 Illustration of dynamic programming, integer programming and Viterbi algorithm. 7
2.2 Methodologies for constrained optimization problem. 7
2.3 Illustration of integer programming. 8
2.4 For each independent source, moving the violated value to boundary value achieves minimal Lagrangian cost based on independent and convex properties. 11
2.5 Illustration of water filling principle. 12
2.6 Illustration of optimization solution by minimal Lagrangian cost. 13
2.7 Different quantizer choice for frame 1 leads to different R-D curve of frame 2, also the solution of minimal lagrangian cost to dependent problem. 14
2.8 Illustration of Viterbi algorithm with minimal Lagrangian cost path in IBBP case. 15
2.9 VA with skip nodes 16
3.1 Illustrative example of the Viterbi algorithm with skip nodes. 21

4.1 Our experiment procedures. 26
4.2 Total MSE(a), coded frames MSE(b) and copy frames MSE(c) in Akiyo when encoding 2 frames. 27
4.3 Total MSE(a), coded frames MSE(b) and copy frames MSE(c) in Mobile when encoding 2 frames. 28
4.4 Total MSE in Akiyo(a) and in Mobile(b) when encoding 7 frames. . . . 28
4.5 Optimal path from low bit budget to high bit budget of (a)Akiyo, (b)Mobile, (c)Football. 29
4.6 Overall distortion may interlace in static sequence like Akiyo(a); the correct Qp selection for each coding frame may not only reduce coding bits but also reduce overall distortion(b); while available bit budget decreases, the optimal solution may choose different frames to be encoded instead of skipping another frame immediately(c). 30
4.7 Dominative lines of different total number of coded frames. 31
4.8 Illustration of heuristic solution for independent problem with multiple constraints by Viterbi algorithm. 32
4.9 Illustration of methodology 1. 33
4.10 Illustration of methodology 2. 33
4.11 Comparison among proposed heuristic solution, methodology 1 and methodology 2 in Akiyo (a), Mobile(b) and Football(c). 34
4.12 Illustration of successive refinement. 34
4.13 Comparison among proposed heuristic solution and its successive refinement(a), methodology 1 and its successive refinement(b), methodology 2 and its successive refinement(c) in Akiyo. 35
4.14 Comparison among proposed heuristic solution and its successive refinement(a), methodology 1 and its successive refinement(b), methodology 2 and its successive refinement(c) in Mobile. 36
4.15 Comparison among proposed heuristic solution and its successive refinement(a), methodology 1 and its successive refinement(b), methodology 2 and its successive refinement(c) in Football. 37
4.16 Proposed algorithm flowchart. At the first stage to decide which frames to be encoded; to allocate bit among selected frames at the second stage. 37

CHAPTER 1

Research Overview

In the beginning of this thesis, we will introduce the blueprint of the proposed quality- and rate-constrained streaming system, including the architecture, the objective and the formulation to our constrained problem; then related works are mentioned. Finally, the contribution and the organization of this thesis are described in the last section.

1.1 Introduction

1.1.1 Architecture

In this section, we propose a quality- and rate-constrained adaptive frame encoding for specific video streaming applications such as surveillance and mobile systems. Assume there exists a system architecture like Figure 1.1, the client requires high quality surveillance video streaming from the server through the internet TCP/IP transmission. The server has a bandwidth estimation mechanism, which estimates average bandwidth BR in every time slot (10~20 second for example) and we can allocate bits among reference and non-reference pictures according to the estimated BR. However, because network congestion occurs from time to time, we need to provide another rate shaping

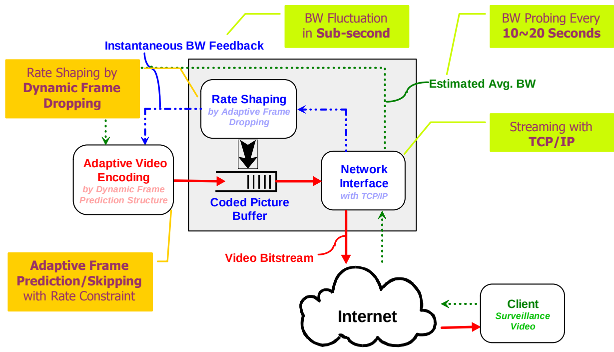


Figure 1.1: Proposed video streaming system architecture.

mechanism such that coded picture buffer can adaptively drop non-reference pictures to adviod short-term bandwidth fluctuation.

Within this architecture, we have constraints and requirements as follows:

- Constraints
  - Real-time and live streaming over interenet
  - Rate- and quality-constrained applications
  - Time-shifted average bandwidth estimation
  - TCP/IP connection
- Requirements
  - A rate- and quality-constrained video encoding scheme
  - A rate shaping mechanism

1.1.2 Objective and Formulation

In order to implement the above system, we need to provide an adaptive frame encoding that

1. Ensures the quality of the reference frames subject to a bitrate constraint.
2. Allows a R-D optimal rate shaping by skipping the non-reference frames.

That is, we have to determine the frames to be skipped from coding and to find the quantization parameters for the frames to be encoded,as Figure 1.2, with the following constraints:

1. The bit of total coded frames is equal to bitrate R_t.
2. The distortion of each coded frame is within [Dmin,Dmax].

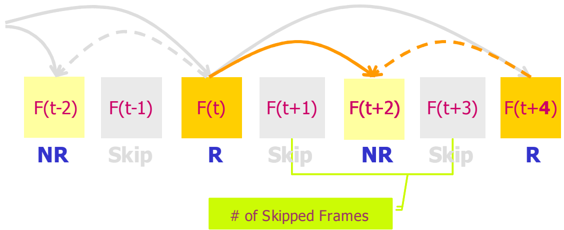


Figure 1.2: Proposed adaptive frame coding.

3. The overall distortion is minimized.

We can further formulate our objectives and constraints as a quality- and rate-constrained optimization problem that

$$\mathbf{q}^* = \arg\min_{\mathbf{q}} \sum_{i=1}^{N} D_i(\mathbf{q})$$

$$s.t. \begin{cases} (1) \sum_{i \in \mathcal{C}} R_i(\mathbf{q}) \leq R_t \\ (2i) \underbrace{D_{\min} \leq D_i(\mathbf{q}) \leq D_{\max}, \forall i \in \mathcal{C}} \end{cases}$$

where

D_i : Distortion of the ith frame

R_i : Rate of the ith frame

$$\mathbf{q} = [q_1, q_2, ..., q_N]^T, q_i = \{0, 1, ..., 51, \underset{Skip}{\boxed{52}}\}, i = \{1, 2, ..., N\},$$

$$\mathcal{C} = \{i : q_i \neq 52, i = \{1, 2, ..., N\}\}$$

Owing to only coded frames having distortion constraints, the number of our constraints is a variable number as $(1 + |\mathcal{C}|)$.

1.2 Related Works

Reed *et al.* used integer programming to analyze the temporal-, spatial- and psnr-domain optimal bit allocation problem under maximal buffer size constraint in [8].

Ortega *et al.* used integer programming as optimal solution to solve the buffer constrained problem for each individual macroblock in [6], also they applied Lagrange multiplier to a nearly optimal solution for the budget constrained problem. Owing to inter programming tests all possible data set to find the optimal solution to multiple constrained or multiple dimensional optimization problems, the solution can be viewed as an absolute optimal solution compared to other algorithms; however, the complexity of integer programming is too large to implement in real time applications. As a result, Lagrangian optimization is applied under several specific assumptions in video coding domain for some fast algorithms.

As for optimization problems with multiple constraints by Lagrangian optimization method, in [5], A. Ortega used Lagrangian method to solve the multiple buffer constrained problem by iteratively adjusting Lagrange multiplier λ. Ahmad *et al.* applied KKT conditions based on Nash bargaining solution and just-noticeable distortion threshold for each macroblock to solving the perceptual quality constrained problem in [1]. In [12], Wang *et al.* also applied KKT conditions to each I frames for the long-term distortion constrained problem. Based on KKT conditions, each constraint corresponds to a Lagrange multiplier and the optimal solution occurs when all costraints are satisfied and the Lagrangian cost is at an minimum value by iteratively adjusting each Lagrange multiplier in video domain.

Furthermore, Ramchandran *et al.* made use of Viterbi algorithm with Lagrangian cost to solve the dependent constrained problem and developed pruning rules based on monotonicity property for the optimization problem with single constraint in [7]. Then Liu *et al.* improved this pruning algorithm for frame skipping situation in [4]. In order to implement adaptive frame skipping in real time system, Song *et al.* predefined Lagrange multiplier and which frames to be encoded for each sub-GOPs, and solved the problem by grandient search in [11]. Because we only have to adjust one Lagrangian multiplier to the optimize problem with single constraint, Viterbi algorithm is applied based on dependent relation and monotonic property. Also, a fast algorithm is developed based on independent relation and commonly used in modern encoder structure.

All constrained problems can be solved by dynamic programming either in constrained form or unconstrained form (Lagrangian method), though the complexity is huge. In order to reduce complexity, modern rate constrained bit allocation problem develop Lagrangian cost algorithm based on independent property for real time applications.

1.3 Contribution and Organization of Thesis

Specially, our main contributions in this work include the following:

- Model for Quality- and Rate-constrained Adaptive Frame Encoding

  We define our adaptive frame encoding problem as a optimization problem and with multiple constraints and survey kinds of methodologies to solve constrained problems.
- Design of Search Strategies for Optimal Solutions

  We implement the dynamic programming to obtain the optimal solution to our quality- and rate- constrained problem and find out the optimal path is a stairway-like curve; besides, we compare the complexity of different dynamic programming algorithms.
- Propose a Heuristic Algorithm Based on Successive Refinement

  We propose a greedy heuristic algorithm based on independency assumption and successive refinement to reduce complexity after observing the optimal solution.

The remaining of this thesis is organized as follows: Chapter 2 contains a survey of constrained optimization problem, and the differences between our problem and other previous works are also compared. Chapter 4 presents the optimal solution and its analysis in the beginning and then we introduce our heuristic solution and the experimental results; also, an iterative algorithm with lower complexity is proposed in the end. This thesis ends with the summary of our observations and a list of future works.

CHAPTER 2

Constrained Optimization Problems : Principles and Applications

---

We will introduce the theory background such as integer programming, Lagrangian optimization and Viterbi algorithm for optimizing constrained problems in detail in the chapter.

2.1 Background

In this section, we will introduce the methodologies to optimize constrained problems and the dynamic programming is commonly used like integer programming and Viterbi algorithm.

A simple algorithm for constrained optimization problem is to find the optimal solution among possible data set by dynamic programming and this algorithm is developed as integer programming; on the other hand, the Lagrange multiplier is another mathematical tool to optimize constrained problems, and the Viterbi algorithm as forward dynamic programming is applied. Although integer programming and Viterbi algorithm are all trellis-based algorithm, the number of nodes at each stage is constant in

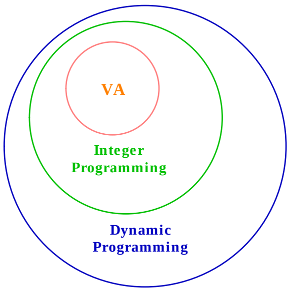


Figure 2.1: Illustration of dynamic programming, integer programming and Viterbi algorithm.

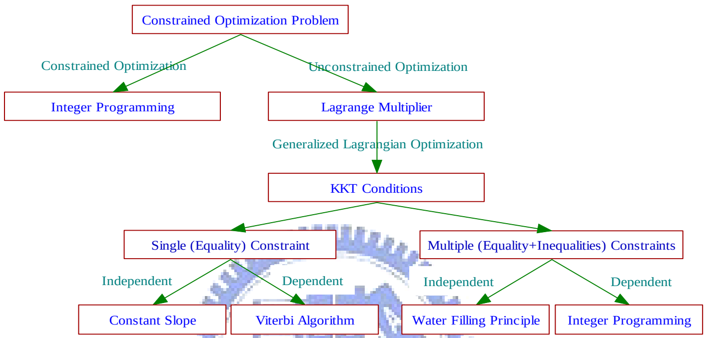


Figure 2.2: Methodologies for constrained optimization problem.

Viterbi algorithm but is variable in integer programming. The relation among dynamic programming, integer programming and Viterbi algorithm is illustrated in Figure 2.1.

Also, Figure 2.2 shows different methodologies to optimize constrained problems, and we will introduce these methodologies in the following subsections: integer programming in section 2.2, Lagrangian optimization in section 2.3, water filling principle in section 2.3.1, and Viterbi algorithm in section 2.3.2.

2.2 Integer Programming

Integer programming is a trellis-based dynamic programming algorithm to solve the constrained problems. Integer programming grows its paths stage by stage and prunes the violated branches to form a trellis from the initial stage to the end stage; as Figure 2.3, the optimal path can be obtained by finding the node with minimal distortion value at the final stage.

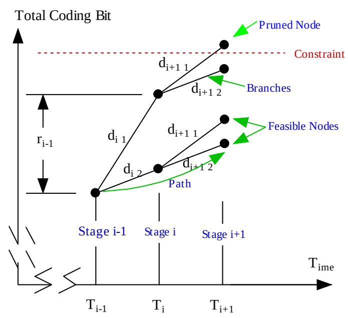


Figure 2.3: Illustration of integer programming.

Although the complexity raising exponentially while the number of stage increases, it can be reduced by clustering the operation points to decrease the number of node or using limited-lookahead window optimization to obtain an sub-optimal path at each window size interval [8].

Ortega *et al.* used the algorithm to solve the problem with buffer constraint on each frame [6]. Additionally, Reed *et al.* analyzed the best solution amaong the combination of spatial, temporal, and psnr dimensions by this algorithm [8].

2.3 Lagrangian Optimization

While dealing with mathematical optimization problems, the method of using Lagrange multiplier λ can be applied to finding the extrema of a function of several variables subject to one or more constraints. That is, suppose a function to be minimized, $f(x, y)$, and the solution set is constrained by another function, $g(x, y) = 0$, the auxiliary function is

$$J(x, y, \lambda) = f(x, y) + \lambda g(x, y)$$

and the minimum occurs when

$$\nabla_{x,y,\lambda} J(x, y, \lambda) = 0$$

Furthermore, owing to the convexity characteristic of the rate-distortion curve in video coding domain, for a given λ, the minimum occurs only at the minimal value of $J(x, y, \lambda) = f(x, y) + \lambda g(x, y)$. As for optmization problems with multiple constraints, the generalized Lagrangian optimization, Karush-Kuhn-Tucker conditions, is commonly used; besides, the water filling principle as a special case of Lagrangian optimization based on independent and convex property is described.

2.3.1 Optimization Problems with Multilple Constraints

2.3.1.1 Karush–Kuhn–Tucker Conditions

The method of using Lagrange multiplier to solve the nonlinear constrained problem is the basic tool in mathematical optimization problems; hence we introduce the generalized Lagrangian optimization (Karush–Kuhn–Tucker Conditions) [3].

Generally, given a optimization problem in the standard form,

$$\min f(\mathbf{x}) \; s.t. \begin{cases} g_i(\mathbf{x}) \leq 0, \;\text{for } i = 1, 2, ..., m \\ h_j(\mathbf{x}) = 0, \;\text{for } j = 1, 2, ..., l \end{cases}$$

where the objective function $f(x)$ is the function to be minimized, and $g_i(x)$, $h_i(x)$ are constraint functions. If x^* is a local minimum, then there exists constants $u_i(i = 1, 2, ..., m)$ and $v_j(j = 1, 2, ..., n)$ such that

$$J(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{m} \mu_i g_i(\mathbf{x}) + \sum_{j=1}^{n} v_j h_j(\mathbf{x})$$

(1) $\nabla_{\mathbf{x}} f(\mathbf{x}^*) + \sum_{i=1}^{m} \mu_i^* \nabla_{\mathbf{x}} g_i(\mathbf{x}^*) + \sum_{j=1}^{n} v_j^* \nabla_{\mathbf{x}} h_j(\mathbf{x}^*) = \mathbf{0}$

(2) $h_j(\mathbf{x}^*) = 0, \;\text{for } j = 1, 2, ..., n$

(3) $g_i(\mathbf{x}^*) \geq 0, \;\text{for } i = 1, 2, ..., m$

(4) $\mu_i^* \geq 0, \text{ for } i = 1, 2, ..., m$

(5) $\mu_i^* g_i(\mathbf{x}^*) = 0, \text{ for } i = 1, 2, ..., m$ (Complementarity)

The above formulation is the famous Karush–Kuhn–Tucker conditions (KKT conditions). It reveals that equalities must set up at the first place and then adjust violated inequalities to boundary values iteratively to obtain minimum x^*. An example is illustrated in section 3.1.

2.3.1.2 Water Filling Principle

Water filling principle is known as a special case of KKT conditions. Suppose the constrained problem to be optimized is

$$\min \sum_{k=1}^{N} D_k(R_k), \ \ s.t. \ \begin{cases} (1) \sum_{k=1}^{N} R_k = RN \\ (2) R_k \geq 0 \end{cases}$$

where R is the average rate. According to KKT conditions, the objective function becomes

$$J(\mathbf{R}, \lambda, \mathbf{u}) = \sum_{k=1}^{N} D_k(R_k) + \lambda(\sum_{k=1}^{N} R_k - RN) + \sum_{k=1}^{N} u_k R_k$$

and optimal solution must satisfy

$$\begin{cases} (1) \nabla J(\mathbf{R}^*, \lambda^*, \mathbf{u}^*) = 0 \Rightarrow \begin{matrix} (1a) D_k'(R_k) = -(\lambda + u_k) \text{ for } k = 1, 2, ..., N \\ (1b) \sum_{k=1}^{N} R_k^* = RN \\ (1c) R_k^* \geq 0 \text{ for } k = 1, 2, ..., N \end{matrix} \\ (4) u_k^* \geq 0 \\ (5) u_k^* R_k^* = 0 \end{cases}$$

Based on the above formulation, the optimal solution can be further discussed into two cases:

***Case* 1** No Violated Inequality

IF $R_k^* > 0$ for all k, then $u_k^* = 0$ for all k (*condition* 5), then the problem degrades to

$$\min \sum_{k=1}^{N} D_k(R_k), \ s.t. \ \sum_{k=1}^{N} R_k = RN$$

and the solution is

$$\underbrace{D_k'(R_k^*) = -\lambda}_{Equal\ slope} \text{ for } k = 1, 2, ..., N, \ \text{ where } \sum_{k=1}^{N} R_k^* = RN$$

which represents the equal slope concept.

***Case* 2** With Violated Inequalities

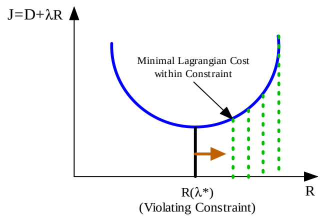


Figure 2.4: For each independent source, moving the violated value to boundary value achieves minimal Lagrangian cost based on independent and convex properties.

The violated inequalities can be adjusted based on the following two assumptions in video domain:

- Independency

  Each source signal is independent of others. As a result, a violated source can be adjusted to boundary value to satisfy the constraint without affecting rate-distortion curves of other sources.

- Convexity

  Each source has one and only one minimal Lagrangian cost operation point based on convexity. Any violated source can be adjusted to boundary value of the constraint to achieve the corresponding constrained local minimum as Figure 2.4.

According to the above assumptions and KKT conditions, if $R_k^* = 0,\ k \in H$ for several source set H, then $u_k^* = 0$ for $k \in T = \{\mathbf{N} - H\}$, then the original problem becomes

$$\min \sum_{k=1}^{N} D_k(R_k),\ \ s.t.\ \begin{cases} (1)\sum_{k=1}^{N} R_k = RN \\ (2) R_k = 0\ for\ k \in H \end{cases}$$

$$\Rightarrow \min \sum_{k \in \boldsymbol{T}}^{N} D_k(R_k),\ \ s.t.\ \sum_{k \in \boldsymbol{T}}^{N} R_k = RN$$

and the solution becomes

$$\underbrace{D_k'(R_k^*) = -\lambda}_{Equal\ slope} \text{ for } k \in \mathbf{N} \backslash \boldsymbol{H},\ \text{ where } \sum_{k \in \boldsymbol{T}}^{N} R_k^* = RN$$

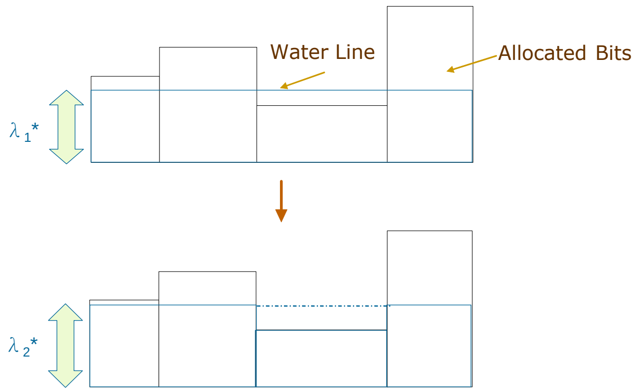


Figure 2.5: Illustration of water filling principle.

That is, the optimal solution includes the following steps:

1. Ingore the inequality constraints and bring constant slope into practice.
2. If all constraint are satisfied then the optimal solution is achieved. Otherwise, adjusting the violated constraints by moveing them to the boundary values and no more optimization operations for them.
3. Update the equality constraint.
4. Repeat step 1, 2, 3 until the optimal solution is achieved or there is no suitable solution for current constrained problem.

A graphic illustration of the solution is shown in Figure 2.5, and it is commonly called a water filling principle.

2.3.2 Optimization Problems with Single Constraint

In [6], a theorem is proposed to solve the budget constrained problem that for any real positive number λ, the Lagrange multiplier, if the mapping $x^*(i)$ for $i = 1, 2, ..., n$ minimizes

$$\sum_{i=1}^{n} d_{ix(i)} + \lambda r_{ix(i)}$$

then it is also the optimal solution to the problem

$$\min \sum_{i=1}^{n} d_{ix(i)},\ s.t. \sum_{i=1}^{n} r_{ix(i)} \leq R_t$$

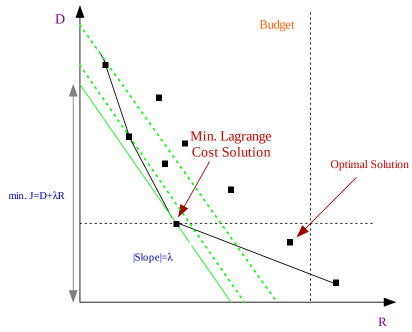


Figure 2.6: Illustration of optimization solution by minimal Lagrangian cost.

where

$$D(\lambda) = \sum_{i=1}^{n} d_{ix^*(i)} \leq \sum_{i=1}^{n} d_{ix(i)},\ s.t. R_t = R(\lambda) = \sum_{i=1}^{n} r_{ix^*(i)}$$

That is, referring to λ as the slope (Figure 2.6), for a fixed λ, we can obtain the best possible solution that meets the budget constraint $R_t = R_{total}$. And the λ is needed to iteratively change by bisection search algorithm [6][10] until we find the multiplier λ^*, such that the total number of used bits meets the original budget constraint, $R(\lambda^*) = R_{total}$, within a convex hull approximation.

Therefore, we can transform the original constrained problem to unconstrained problem, and the solution by this constant slope algorithm is optimal for rate distortion trade-off. For example, the typical rate control problem is defined to minimize the total distortion $\sum_{i=1}^{n} D_i(Q_1, Q_2, ..., Q_i)$, where $Q_i \in \{q_1, q_2, ..., q_N\}$, subject to the total rate/budget constraint R_{total} as follows:

$$\min \sum_{i=1}^{n} D_i(Q_1, Q_2, ..., Q_i),\ \ s.t.\ \sum_{i=1}^{n} R_i(Q_1, Q_2, ..., Q_i) \leq R_{total}$$

and the optimal solution is equal to finding Q^*, λ^* and to minimize

$$J(Q, \lambda) = \sum_{i=1}^{n} J_i(Q_1, Q_2, ..., Q_i),$$

$$where\ J_i(Q_1, Q_2, ..., Q_i) = D_i(Q_1, Q_2, ..., Q_i) + \lambda R_i(Q_1, Q_2, ..., Q_i)$$

such that $J(Q^*, \lambda^*) \leq J(Q, \lambda^*)$ and $R(\lambda^*) = R_{total}$.

Owing to that current coding unit can reference previous coded units to reduce tem-

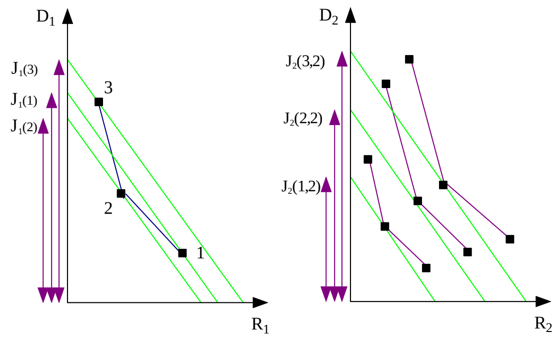


Figure 2.7: Different quantizer choice for frame 1 leads to different R-D curve of frame 2, also the solution of minimal lagrangian cost to dependent problem.

poral redundancy in video coding, the quality of previous coded units will impacts the coding efficiency of the following coding units. That is, the sum of minimal Lagrangian cost at each individual stage will not always result in the optimal solution, as Figure 2.7 [7]. Song *et al.* assumed the dependency relationship between sub-GOPs and propose a real-time system to optimize low-bitrate constrained problem with frame skipping, where the Lagrange multiplier and frames to be encoded for each sub-GOPs are pre-decided [11]. Schuster *et al.* also developed an MINMAX distortion criterion based on Lagrangian method to solve the minimum rate subject to each source distortion constrained dependent problem [9].

In order to solve this dependent problem, the Viterbi algorithm with Lagrangian cost is applied and we will introduce it in section 2.3.2.1.

2.3.2.1 Viterbi Algorithm for Dependent Problems

Viterbi algorithm [2] is a trellis-based forward dynamic programming procedure which iteratively determines possible shortest paths and prunes out non-optimal paths stage by stage.

For each stage, a node is an operating point of a quantizer, and a growing branch is connected from node at the previous stage to node at the current stage with corresponding Lagrangian cost $J = D + \lambda R$. The optimal solution is the path of minimal Lagrangian cost from the beginning stage to the end stage for a specific λ. When λ increases, the optimal path is tend to smaller the total coding bits, and vice versa. Still, we have to iteratively find λ^* by bisection search until $R(\lambda^*) = R_{total}$.

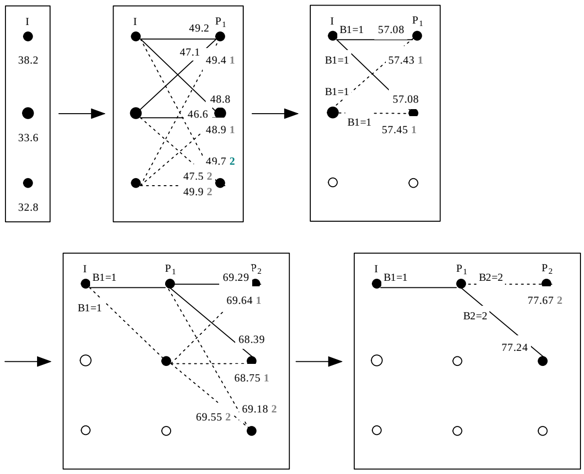


Figure 2.8: Illustration of Viterbi algorithm with minimal Lagrangian cost path in IBBP case.

Ramchandran *et al.* applied VA algorithm with Lagrangian cost and propose pruning rules based on monotonicity property, as Figure 2.8 [7]. Assume the quantizer grades ordered from finest to coarsest, for any $\lambda \geq 0$, there exists monotonicity property that for $i \leq i'$

$$J_2(i,j) < J_2(i',j)$$

where quantizer j of frame 2 is dependent on quantizer i of frame 1. Afterwards, the pruning conditions based on monotonicity property are used to eliminate suboptimal operating points:

- Pruning Condition 1

  $If\ J_1(i) + J_2(i,j) < J_1(i') + J_2(i',j)$, for any $i < i'$, then (i',j) can be pruned out.

- Pruning Condition 2

  $If\ J_2(i,j) < J_2(i,j')$, for any $j < j'$, then (i,j') can be pruned out.

- Pruning Condition 3

  According to monotonicity property and pruning conditions, if $J_1(i) < J_1(i')$ for $i < i'$, then state node i' can be pruned (for I frames).

In [4], Liu *et al.* also improved VA by considering frame skipping situation, as Figure 2.9 [4], and assumed monotonicity property with frame skipping brings into being for

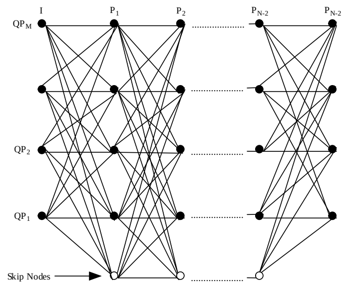


Figure 2.9: VA with skip nodes

any $\lambda \geq 0$,

$$J(i, s_{ij}, j) \leq J(i', s_{i'j}, j), \; if \; i \leq i'$$
$$J(i, s_{ij}, j) \leq J(i, s_{ij'}, j'), \; if \; j \leq j'$$
$$J(i, s_{ij}, j) \leq J(i', s_{i'j'}, j'), \; if \; i \leq i', \; j \leq j'$$

where s_{ij} represents skipped frame reconstructed from forward coding frame with quantizer i and backward coding frame with quantizer j. Also, the new pruning rules are:

- Pruning Condition 4

  $If \; J(i) + J(i, s_{ij}, j) + J(i, j) \leq J(i') + J(i', s_{i'j}, j) + J(i', j)$, for $i < i'$ then branch $J(i', s_{i'j}, j)$ can be pruned out.
- Pruning Condition 5

  $If \; J(i, s_{ij}, j) + J(i, j) \leq J(i, s_{ij'}, j') + J(i, j')$, for $j < j'$ then branch $J(i, s_{ij'}, j')$ can be pruned out.

2.3.2.2 Viterbi Algorithm for Independent Problems

Though dynamic programming can apply to finding the optimal solution to dependent problems, it takes too much computation considering frame-to-frame dependency. In order to reduce complexity, solution to dependent problem usually reduced to independent problem; that is, take rate constrained problem for example, the problem formulation becomes

$$\min \sum_{i=1}^{n} D_i(Q_i), \; s.t. \; \sum_{i=1}^{n} R_i(Q_i) \leq R_{total}$$

The solution to independent problem only focuses on finding the minimal Lagrangian cost at current stage despite of the effect of other stages.

2.4 Complexity Comparison

Table 2.1 lists the complexity of two different dynamic programming procedures, integer programming and Lagrangian cost based Viterbi algorithm, for the optimal solution to optimization problems with multiple constraints.

In the dependent case, these two DP algorithms both grow exponentially. However, in the independent case, Lagrange multiplier method can reduce complexity enormously by selecting the minimal Lagrangian cost node at each stage, though it still has to iteratively refine the value of Lagrange multiplier; as for integer programming, because it does not have a independent form, the complexity remains the same.

Besides, owing to each constraint corresponding to a Lagrange multiplier, and the optimal solution by Lagrangian method needs to solve each Lagrange multiplier iteratively, integer programming for multiple constrained case is better than Lagrangian optimization for lower complexity; however, as for singular equality constrained problem, Lagrangian optimization is more suitable for only one Lagrange multiplier to be solved, and a fast algorithm can be developed based on independency assumption. The following lists the definition of variables:

N : Number of Coding Units
M : Number of Operating Modes per Coding Units
C : Number of Constraints
T : Number of Test Points per Lagrange Multiplier

Table 2.1: Complexity comparison between DP

Integer Programming	**Lagrange Multiplier**
Optimal	Convex Approx.
N^M (Dependent)	$T^C N^M$ (Dependent)
N^M (Independent)	$T^C MN$ (Dependent)

CHAPTER 3

Related Works

In this chapter, we will introduce how related works applied the optimization methods described in previous chapter into their constrained problem. Also, the difference between related works and our proposed problem is listed in the end.

3.1 KKT Conditions

Wang *et al.* used Karush–Kuhn–Tucker Conditions to solve long-term distortion constrained problem for I frames [12]; the fomulation and algorithm for their problem are in the following:

- ***Problem Formulation***

  Let ϑ be a set of quantizers and let $D_{\min}$ and $D_{\max}$ be the lower and upper bounds of the distortion for each source sample. Find $Q^* = (Q_1^*, Q_2^*, ..., Q_n^*)$, with $Q_i^* \in \vartheta$ for $i = 1, 2, ..., n$, where n is the number of source samples, such that

$$Q^* = \arg\min_{Q\in\vartheta} \sum_{i=1}^{n} R_i(Q_i)$$

$$s.t.\begin{cases}(1)D_{\min} \le D_i(Q_i) \le D_{\max}, i = 1, 2, ..., n \\ (2)\sum_{i=1}^{n} D_i(Q_i) \le D_{total} = \frac{(D_{\min}+D_{\max})}{2} \cdot n\end{cases}$$

- ***Algorithm***

1. Consider the original optimization problem with multiple constraints as an equivalent problem with a total "distortion budget", $D_{total} = \frac{(D_{\min}+D_{\max})}{2} \cdot n$.
2. Apply the Lagrangian method to solve the problem with $D_{\min} = 0$ and $D_{\min} = \infty$. The result is the constant slope solution with optimal λ^0 and corresponding Q^0, such that $\sum_{i=1}^{n} D_i(Q_i) \approx D_{total}$. The approximation is due to the fact that the operational rate-distortion function is a discrete function.
3. Impose the distortion constraints. The quantization found in the previous step is the optimal solution that minimizes total bit rate for a given total distortion. For any frame i, the constraint condition $D_{\min} \le D_i(Q_i^0) \le D_{\max}$ may be violated. Depending on the value of $D_i(Q_i^0)$, frames are divided into three groups:
   (a) If $D_i(Q_i^0) < D_{\min}$, the constant slope solution for this frame is not admissible. In this case, we need to replace Q_i^0 by an admissible Q_i^* such that $D_i(Q_i^*) \approx D_{\min}$.
   (b) Similarly, if $D_i(Q_i^0) \ge D_{\max}$, replace Q_i^0 by an admissible Q_i^* such that $D_i(Q_i^*) \approx D_{\max}$.
   (c) If $D_{\min} \le D_i(Q_i^0) \le D_{\max}$, the constant slope solution $D_i(Q_i^0)$ does not violate the distortion constraints, but, due to the changes of the operating points of the frames in the other two groups, the Q values for this group cannot be finalized at this stage.
4. Initialize the next iteration. Let the number of frames in the above three groups be $N_{\min}$, $N_{\max}$, and N_{mid}, respectively. If $N_{\min} = N_{\max} = 0$, then the constant slope solution found in Step 2 is also the solution to the given distortion-constrained problem. Otherwise, performs

$$N \leftarrow N_{mid}$$

$$D_{total} \leftarrow D_{total} - N_{\min}D_{\min} - N_{\max}D_{\max}$$

   If the updated D_{total} value is positive, go back to Step 2; otherwise, end the algorithm and the Q found at this iteration is the approximated solution to the given distortion-constrained problem.

3.2 Viterbi Algorithm with Lagrangian Cost

Liu *et al.* applied Viterbi algorithm with Lagrangin cost to solving the spatial quality (QP) selection, temporal resolution (frame rate) optimization problem in [4], where they use MCI as a temporal interpolation method to reconstruct skipped frames. The formulation and algorithm for their problem are in the following:

- ***Problem Formulation***

  Let ϑ be a set of quantizers ranging from $Q_{\min}$ to $Q_{\max}$. Find $Q^* = (Q_1^*, Q_2^*, ..., Q_n^*)$, with $Q_i^* \in \vartheta$ for $i = 1, 2, ..., n$, where n is the number of source samples, and find $S^* = (S_1^*, S_2^*, ..., S_n^*)$ with $S_i \in [0, 1]$, where $S_i = 0$ represents current frame is skipped, $S_i = 1$ represents current frame is encoded; the maximal number of successive skipping frame $S_{\max} = 2$, such that

$$(Q^*, S^*) = \arg \min_{Q \in \vartheta, S \in [0,1]} \sum_{i=1}^{N} D_i(Q, S) = \sum_{i=1}^{N} \{D_i(Q_i)|(S_i = 1) + D_i(Q_i)|(S_i = 0)\}$$

$$s.t. \sum_{i=1}^{N} R_i(Q, S) = \sum_{i=1}^{N} R_i(Q_i)|(S_i = 1) \leq R_{budget}$$

where

$$S = [S_1, S_2, ..., S_N], S_i \in [0, 1], i = 1, ..., N$$
$$\vartheta = [Q_1, Q_2, ..., Q_N], Q_i \in [Q_{\min}, Q_{\max}], i = 1, ..., N$$

- ***Algorithm***

1. Initialize the value of λ.
2. Calculate for the first frame, which is an I-frame, for every QP value within the range $i \in [Q_{\min}, Q_{\max}]$, as shown in Figure 3.1 (a).
3. Prune unqualified I-nodes according to the monotonicity property as shown in Figure 3.1 (b).
4. Grow the trellis to Stage 2 by coding the first P-frame with all QP values. The skip node is reserved as shown in Figure 3.1 (c).
5. Prune at Stage 2 with Rules 1 and 2. Note that the skip node should be kept as

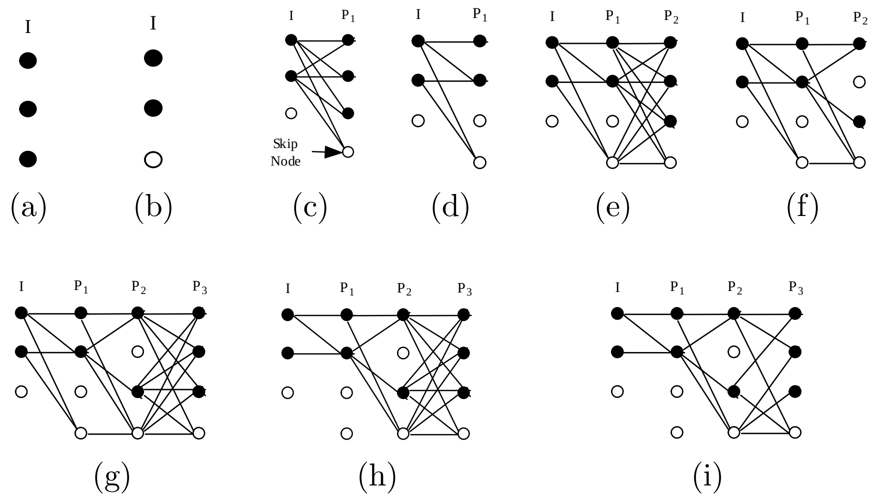


Figure 3.1: Illustrative example of the Viterbi algorithm with skip nodes.

shown in Figure 3.1 (d).

6. Grow the trellis to one more stage. The skipped frame in the previous stage is reconstructed by the neighboring reference frames coded with selected QPs as shown in Figure 3.1 (e).
7. Prune at Stage 3 based on the monotonicity property, i.e., Rules 1 and 2 for pruning the third coded frames, Rules 3 and 4 for pruning previous skipped frames. The skip node at Stage 3 is reserved as shown in Figure 3.1 (f).
8. Similar to Step 4, grow trellis to Stage 4 as shown in Figure 3.1 (g).
9. Prune paths that have more successive skipped frames than $S_{\max}$ as shown in Figure 3.1 (h).
10. Similarly to Step 7, pruning is performed based on the monotonicity property as shown in Figure 3.1 (i).
11. Repeat Step 8–10 until the last frame. Update and return to Step 2.
12. Stop when λ converges.

3.3 Variable Frame Rates Encoding

Song et al. proposed a rate control mechanism for low-bit-rate video via variable-encoding frame rates in [4]. In order to implement this variable frame rate encoding under real time environment, they divide each gop into 8 sub-GOPs with size=12 for complexity issue and define the value of each Lagrange multiplier for each sub-GOP under the consideration of dependency relationship between sub-GOPs. The proposed algorithm consist of two parts:

1. Frame-Rate Control

   Decide the frame rate (number of encoded frames) and encoded frame positions for each sub-GOP before encoding based on the histogram of difference image.

2. Bit Allocation

   Use gradient search method to obtain the optimal QP setting for individual frame encoding in each sub-GOP, such that the following problem statement is satisfied.

- ***Problem Formulation***

  Determine $\vec{q_m}$, m=1,2,...,M to minimize

$$\sum_{m=1}^{M}(D_m(\vec{q_m}) + w_q E_m(\vec{q_m}))$$

$$s.t. \sum_{m=1}^{M} r_m(\vec{q_m}) \leq B_{subgop} \cdot M$$

where $\vec{q_m} = (q_{m,1}, q_{m,2}, ..., q_{m,N_m})$ is the quantization parameter vector for the mth sub-GOP, and

N_m	:	encoded frame number of the mth sub-GOP;
$r_m(\vec{q_m})$	:	assigned number of bits for the mth sub-GOP;
M	:	number of sub-GOPs in a GOP;
N_{subgop}	:	total frame number of a sub-GOP;
N_{gop}	:	total frame number of a GOP;
w_q	:	weighting factor for abrupt quality change and flickering;
$d_i(q_1, q_2, ..., q_i)$	:	distortion measure for the ith frame;
$r_i(q_1, q_2, ..., q_i)$	:	allocated bit rates for the ith frame;

and

$$D_m(\vec{q_m}) = \frac{1}{N_m}\sum_{i=1}^{N_m} d_i(q_1, q_2, ..., q_i),$$

$$E_m(\vec{q_m}) = \frac{1}{N_m}\sum_{i=1}^{N_m} (d_i(q_1, q_2, ..., q_i) - d_{i-1}(q_1, q_2, ..., q_{i-1}))^2$$

$D_m(\vec{q_m})$ represents the distortion measure of encoded frames, and $E_m(\vec{q_m})$ represents the distortion variance measure of two successive encoded frames; besides, w_q is a weighting factor for abrupt quality change and flickering controlling and it is set to 2

Table 3.1: Comparison among the proposed problem and related works

Author	Thesis	Wang *et al.* [12]	A. Ortega [5]
Constraints	Rate, Distortion(Multiple)	Distortion(Multiple)	Buffer(Multiple)
Frame Skipping	Y	N	N
Recovery Method	Frame Copy	NA	NA
Frames to be Encoded	Unknown	Pre-Decided	Pre-Decided
Lagrange Multiplier Value	Bisection Search	Bisection Search	Bisection Search
Proposed Algorithm	VA	KKT Conditions	KKT Conditions
Author		Schuster *et al.*[9]	Ahmad *et al.*[1]
Constraints		Distortion(Multiple)	Perceptual Quality(Multiple)
Frame Skipping		N	N
Recovery Method		NA	NA
Frames to be Encoded		Pre-Decided	Pre-Decided
Lagrange Multiplier Value		Bisection Search ($D_{\max}$)	Bisection Search
Proposed Algorithm		DP	KKT Conditions
Author		Liu *et al.* [7]	Song *et al.* [11]
Constraints		Rate	Rate
Frame Skipping		Y	Y
Recovery Method		Temporal Interpolation	Frame Copy
Frames to be Encoded		Unknown	Pre-Decided
Lagrange Multiplier Value		Bisection Search	Pre-Decided
Proposed Algorithm		VA	Gradient Search

in their experiments.

After defining the penalty function for the mth sub-GOP as

$$P_m(\vec{q_m}) = \sum_{i=1}^{m} r_i(\vec{q_i}) - m \cdot B_{subgop}$$

and applying the pre-defined Lagrange multiplier λ_m, the above constrained problem becomes to minimize the following unconstrained function

$$\Phi_m(\vec{q_m}, \lambda_m) = J_m(\vec{q_m}) + \lambda_m \max\{0, P_m(\vec{q_m})\}, \ for\ m = 1, 2, ..., M$$
$$where\ J_m(\vec{q_m}) = D_m(\vec{q_m}) + w_q E_m(\vec{q_m})$$

and a gradient search method was used to find the optimal solution.

3.4 Comparison

In order to understand the differences between the proposed quality- and rate-constrained problem and other constrained problems, we compare and list several features in Table 3.1.

From the above table, we can clearly observe that our problem is different from the others and our problem is difficult for having variable multiple constraints, uncertain to number of skipped frames, and the dependent relation between frames to be skipped and frames to be encoded. Besides, The best method of getting the optimal solution to our constrained problem is to use integer programming instead of Lagrangian optimization.

CHAPTER 4

Experiments and Analyses

In this chapter, we implement the integer programming to solve our quality- and rate-constrained problem, and analyze the optimal solution path; then, in order to reduce complexity, we propose a simple heuristic solution based on independent relationship and successive refinement assumption. Figure 4.1 shows our experiment procedures; the experiment settings, experiment results and observations for each experiment are described in each section. We use jm 12.3 and CIF format with 37~39 psnr constraints in all experiments.

4.1 Optimal Solution Analyses

We implement the integer programming algorithm to solve our constrained problem, and observe the R-D data from the view of MSE, optimal path, and then propose the concept of dominant lines.

4.1.1 MSE Weighting Effect

Setting:

Sequence	Akiyo, Mobile
Algorithm	Integer Programming
Total Frame	8
Experimental Group	Coded Frame MSE, Copy Frame MSE, Total MSE with Frames to be Encoded=7
Control Group	Coded Frame MSE, Copy Frame MSE, Total MSE with Frames to be Encoded=2

Observation

- Copy frame MSE effect:

  From Figure 4.2 and 4.3, we can clearly observe the copy frame MSE weights overwhelmingly compared to the coded frame MSE.

- Variation line in local view, horizontal line in global view:

  Though the coded frame MSE varies among different coding bit range in Figure 4.2 (b) and 4.3 (b), the overall MSE looks like horizontal lines for copy frame MSE effect.

- Larger MSE gap with more frame skipped:

  From Figure 4.2 (a), 4.4 (a) and Figure 4.3 (a) ,4.4 (b), the gap between different coding frame selections is larger while more frames are skipped.

- Unobvious dominant lines in static sequence:

  There does not exist obvious horizontal dominant lines in static sequence, however, the gap among different coding frame selections is minor with less frame

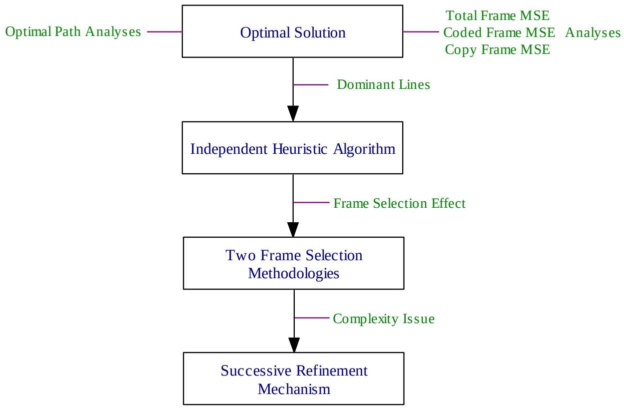
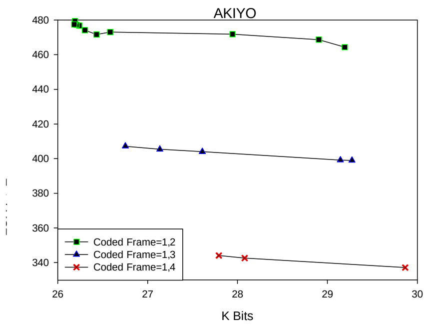
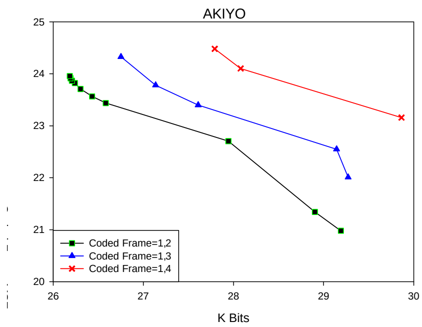


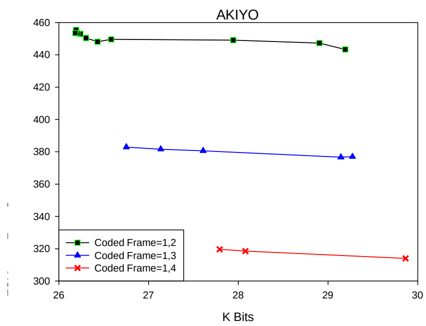
Figure 4.1: Our experiment procedures.



(a)



(b)



(c)

Figure 4.2: Total MSE(a), coded frames MSE(b) and copy frames MSE(c) in Akiyo when encoding 2 frames.

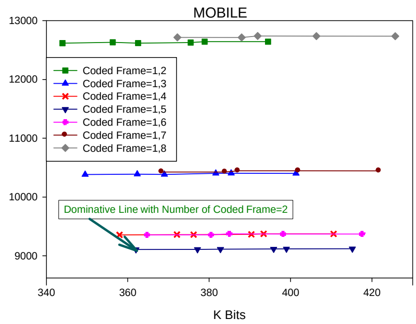
skipped as Figure 4.2 (a) and 4.4 (a).

- Obvious dominant lines in motion sequence:

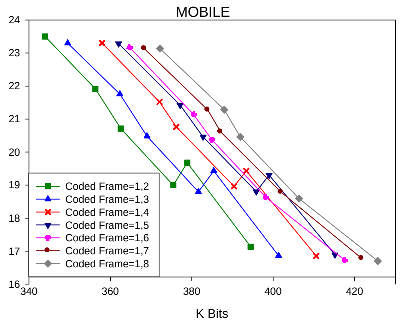
  There exists a dominant line for each different total coding frame in motion sequence as Figure 4.3 (a) and 4.4 (b).

Remarks

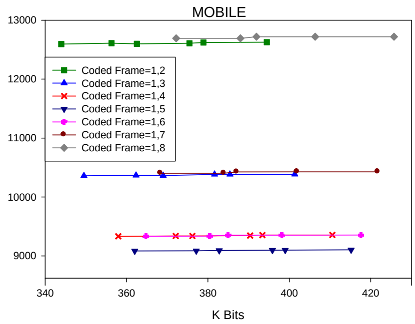
Because the copy frame MSE weights overwhelminly compared to coded frame MSE, choosing the correct combination of coding frames is more important than allocating bits among these frames; besides, since the gap between different coding frame selections is large and there exists a dominant line for each different total coding frames, we can reduce the original multiple constraints dependent problem to independent problem by always finding the corresponding dominant line of current encoding frame numbers, which is the basic idea of our heuristic solution.



(a)

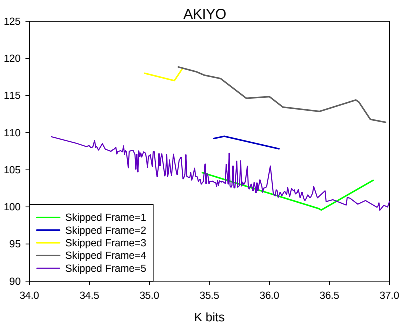


(b)

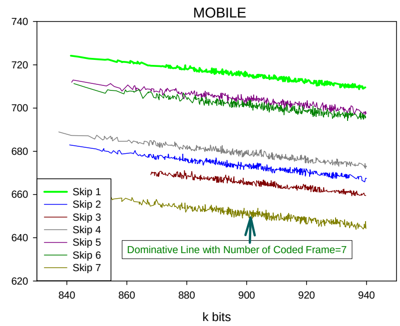


(c)

Figure 4.3: Total MSE(a), coded frames MSE(b) and copy frames MSE(c) in Mobile when encoding 2 frames.

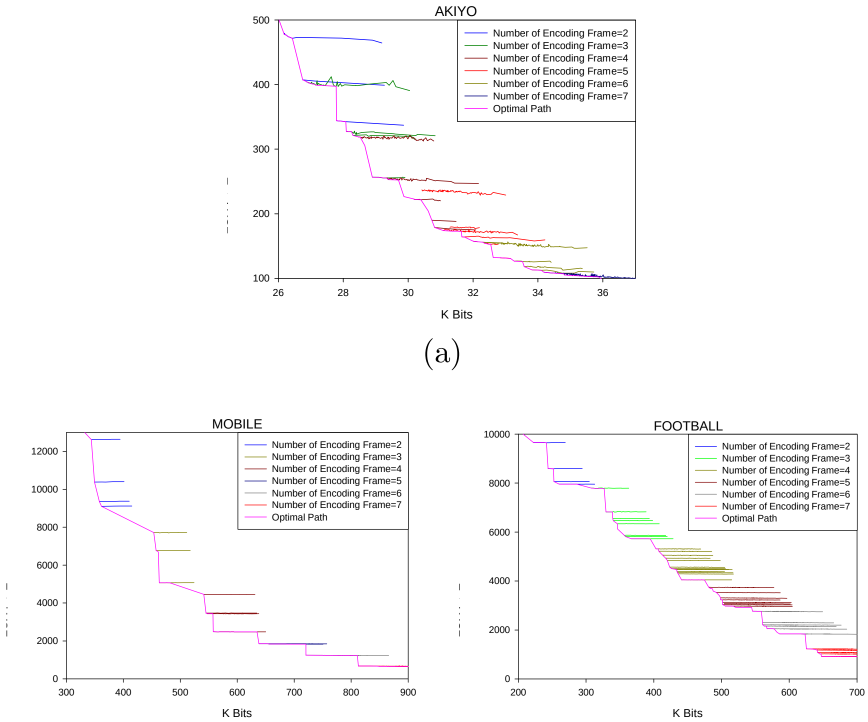


(a)



(b)

Figure 4.4: Total MSE in Akiyo(a) and in Mobile(b) when encoding 7 frames.



(a)

(b) (c)

Figure 4.5: Optimal path from low bit budget to high bit budget of (a)Akiyo, (b)Mobile, (c)Football.

4.1.2 Optimal Path Analyses

Setting:

Sequence	Akiyo, Mobile, Football
Algorithm	Integer Programming
Total Frame	8

Observation

- Optimal path is a stairway-like curve:

  Figure 4.5 shows the optimal path of (a)Akiyo, (b) Mobile, (c) Football when coding bits range from low to high. The optimal path is a stairway-like curve, which means number of frames to be encoded and coding frame selection should be different to optimize our constrained problem while the available bit budget changes.

- Optimal path chooses different total coding frames interactively in static sequence:

  In static sequence like Akiyo, when available bit budget decreases, the number

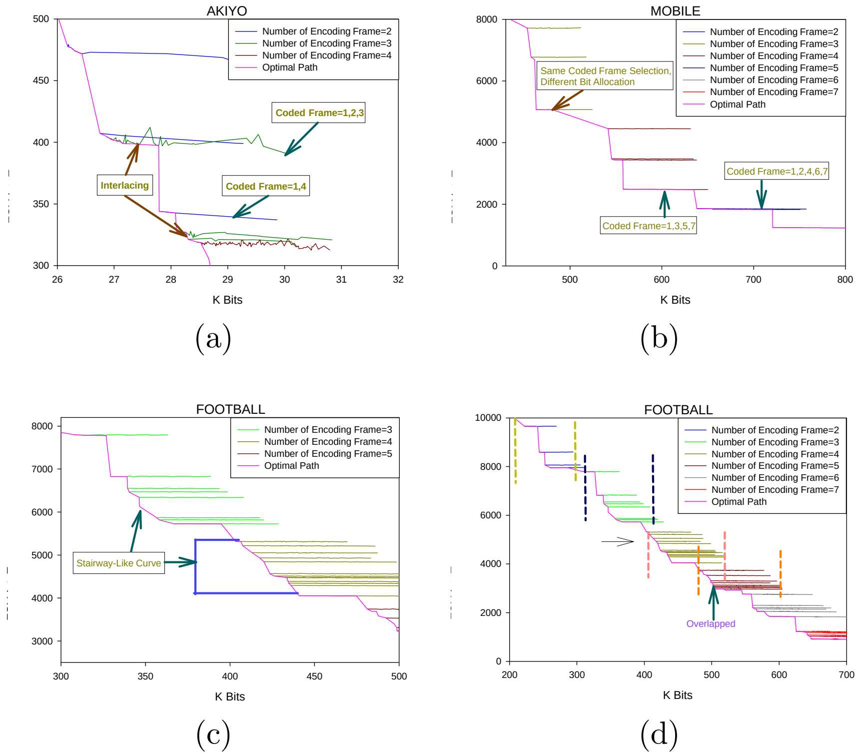


Figure 4.6: Overall distortion may interlace in static sequence like Akiyo(a); the correct Qp selection for each coding frame may not only reduce coding bits but also reduce overall distortion(b); while available bit budget decreases, the optimal solution may choose different frames to be encoded instead of skipping another frame immediately(c).

of total coding frames may increase, as the interlacing part in Figure 4.6 (a); however, this situation would not occur in motion sequence such as Football and Mobile for Frame copy MSE weights largely.

- Encoding more bits does not always results in smaller distortion:

  Not always using more bits results in smaller distortion, as Figure 4.6 (b). Because skipped frame makes temporal dependency decrease and the frame copy effect, even we allocate bits among the same selected frames, the correct Qp setting for each frame sometimes not only reduce coding bits but also reduce overall distortion.

- Choosing another coding frame combination is better than skipping another frame immediately:

  In Figure 4.6 (c) we can observe the clear stair-way curve, which means the optimal path should select another frame combination instead of choosing another frame to be skipped immediate when bit budget decrease.

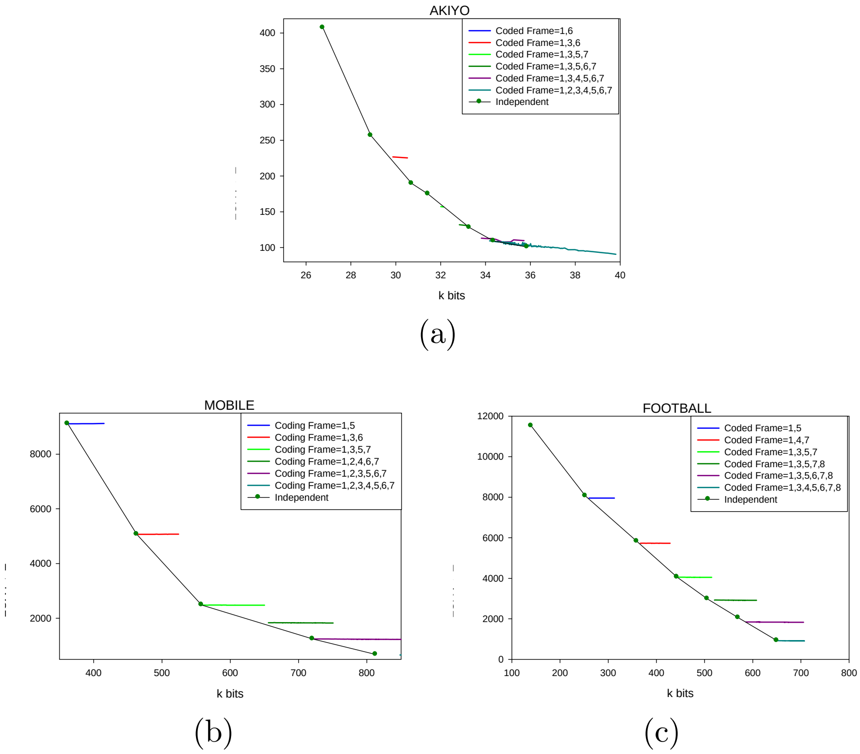


Figure 4.7: Dominative lines of different total number of coded frames.

- Selecting another frame to be skipped immediately for complexity issue when less frame skipped:

  From Figure 4.6 (d), the bit range of number of coded frame=5 is overlapped with the bit range of number of coded frame=4; however, the bit range of number of coded frame=3 is not overlapped with the bit range of number of coded frame=2. That is, when available bit decreases, to select a frame to be skipped is a good choice with less frame skipped, but it is better to select another coding frame combination with more frame skipped.

Remarks

The optimal solution to our proposed constrained problem is to use dynamic programming. However, in Figure 4.5, each line represents a coding frame combination, e.x. when enocding 3 frames, 1-4-5, among total 8 frames, the optimal solution chooses different coding frame combinations to adapt different bit budget constraint and it must be iteratively solved. Therefore, we want to reduce complexity from dependent problem to independent problem, which becomes to find each dominate line among different total number of coded frames as Figure 4.7.

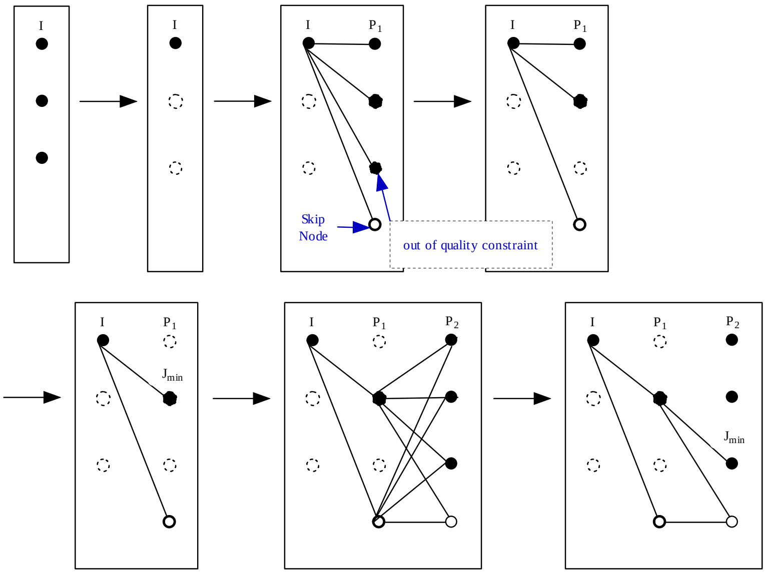


Figure 4.8: Illustration of heuristic solution for independent problem with multiple constraints by Viterbi algorithm.

4.2 Proposed Heuristic Solution

Figure 4.8 illustrates the algorithm flow, and described in the following:

1. Set $\lambda = 0$.
2. Choose the quantization parameter operation point of I frame where lagrangian cost $J = D + \lambda R$ is minimal.
3. Grow the branch from the minimal node of this stage to next stage with $QP_{next} = QP_{current} \pm 5$, and prune out the node with unsatisfied quality constraint.
4. Select the operation point with minimal lagrangian cost among current stage, and repeat 3, 4 until to the end stage.
5. Increase λ gradually until any more increase to λ will result in skipping another frame.

4.2.1 Coding Frame Selection Methodologies

Setting:

Sequence	Akiyo, Mobile, Football
Total Frame	8
Experimental Group	Two Coding Frame Selection Methodology with Heuristic Algorithm
Control Group	Non-Coding Frame Selection Methodology with Heuristic Algorithm

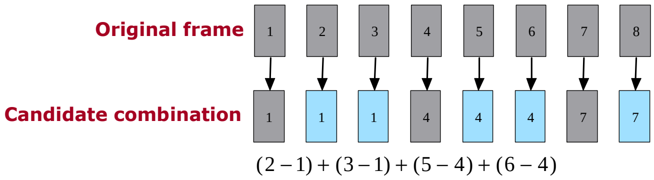


Figure 4.9: Illustration of methodology 1.

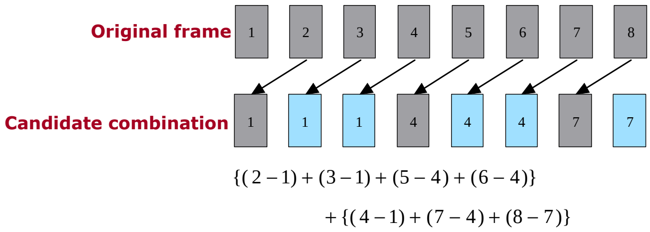


Figure 4.10: Illustration of methodology 2.

The result is in Figure 4.11

1. Suppose there are total i frames in the pre-analysis buffer, and j skipped frames to be chosen, the overall candidates are $\frac{i!}{i!(i-j)!}$ combinations. This methodology is towards to find the minimal sum of difference between coding frame and copy frame. As Figure 4.9, suppose there are 8 frames in the buffer, and we are going to choose frame 1, 4, and 7 as coding frames.
2. Almost the same as methodology 1, excepts that this methodology also considers the difference between coding frames. Figure 4.10 shows a candidate combination.

Observation

- In static sequence, methodology 1 and 2 doesn't have too much performance difference.
- Methodology 1 obtains smaller MSE curve than methodology 2 in motion sequence.
- These two methodology doesn't perform too bad compared with independent R-D curve.

Remarks

Once we know how many coded frames available to encode, we can use these two methodologies to choose which frames to be encoded and reduce complexity compared

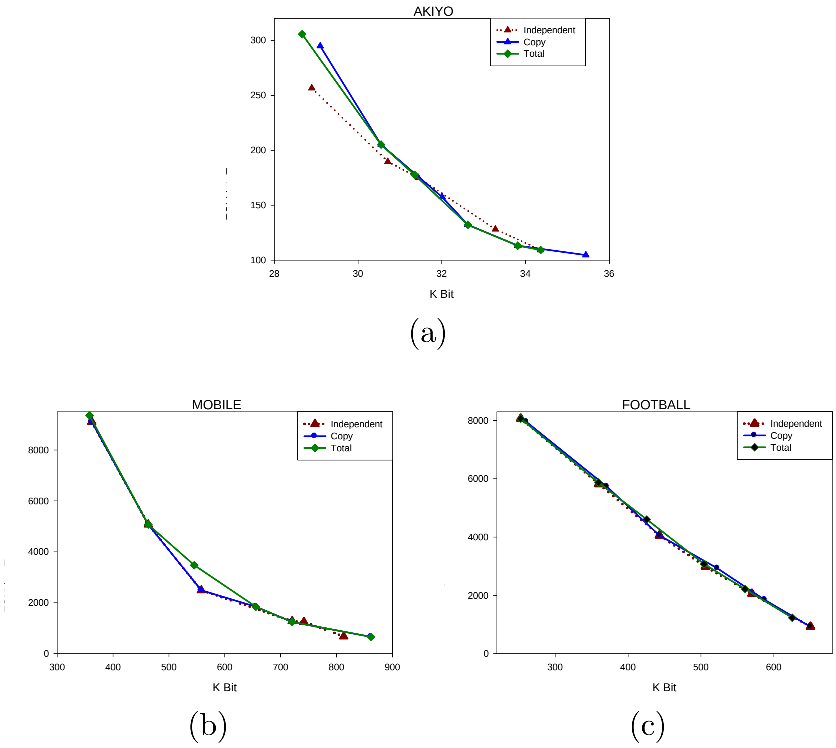


Figure 4.11: Comparison among proposed heuristic solution, methodology 1 and methodology 2 in Akiyo (a), Mobile(b) and Football(c).

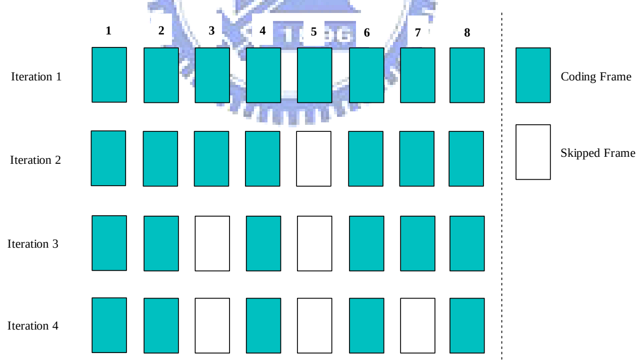


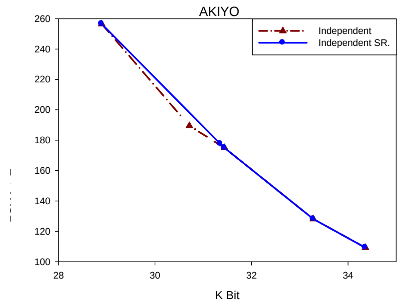
Figure 4.12: Illustration of successive refinement.

to independent solution, though it still needs to try $\frac{i!}{i!(i-j)!}$ combinations, which means once the window size increases, the calculation will also increases exponentially.

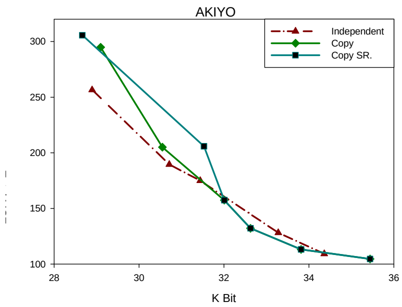
4.2.2 Successive Refinement

Definition of Successive Refinement

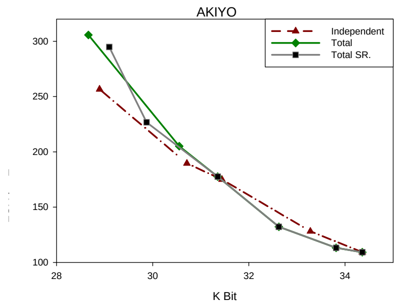
Once we decide to skip frame i, then when it is necessary to skip another frame,



(a)



(b)



(c)

Figure 4.13: Comparison among proposed heuristic solution and its successive refinement(a), methodology 1 and its successive refinement(b), methodology 2 and its successive refinement(c) in Akiyo.

then frame i must be skipped first, as Figure 4.12.

Setting:

Sequence	Akiyo, Mobile, Football
Total Frame	8
Experimental Group	Successive Refinement
Control Group	Non-Successive Refinement

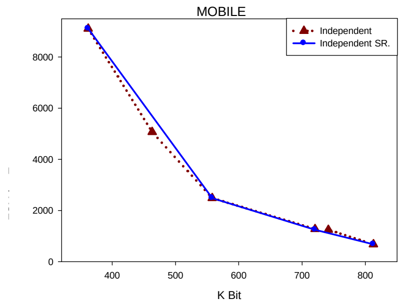
The result for Akiyo is in Figure 4.13, for Mobile in Figure 4.14 and for Football in Figure 4.15

Observation

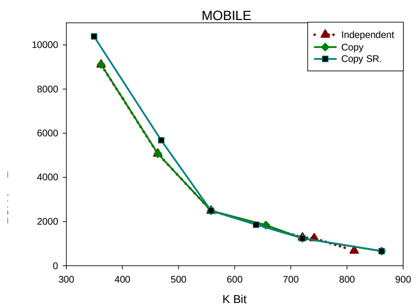
No matter which coding frame selection methodology is used, the difference between successive refinement R-D curve and non-successive refinement R-D curve is minor.

Remarks

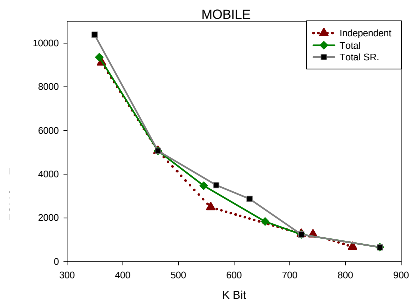
While using successive refinement, the complexity can reduce from exponential increasing to linear increasing.



(a)



(b)



(c)

Figure 4.14: Comparison among proposed heuristic solution and its successive refinement(a), methodology 1 and its successive refinement(b), methodology 2 and its successive refinement(c) in Mobile.

4.3 Proposed Rate Control Mechanism

As Figure 4.16, we propose an iterative successive refinement mechanism to solve the original problem in the independent way. The main idea of the algorithm is to decide which frames to be encoded at the first place for MSE of different copy frames combinations are huge; then we can allocate available bits among these chosen frames to encode and achieve the objective of total minimal distortion. Suppose that there is a R-D model which can estimate the needed coding bits, the rate control algorithm is the following:

Step 1 :

Use the r-d model to calculate the needed coding bits while encoding pre-selected frames at distortion=D_{max}. If the required bits are smaller than available bits, go to Step 3, else go to Step 2.

Step 2 :

Use frame selection methodology to select a frame to be skipped with successive refinement mechanism, and go to Step 1.

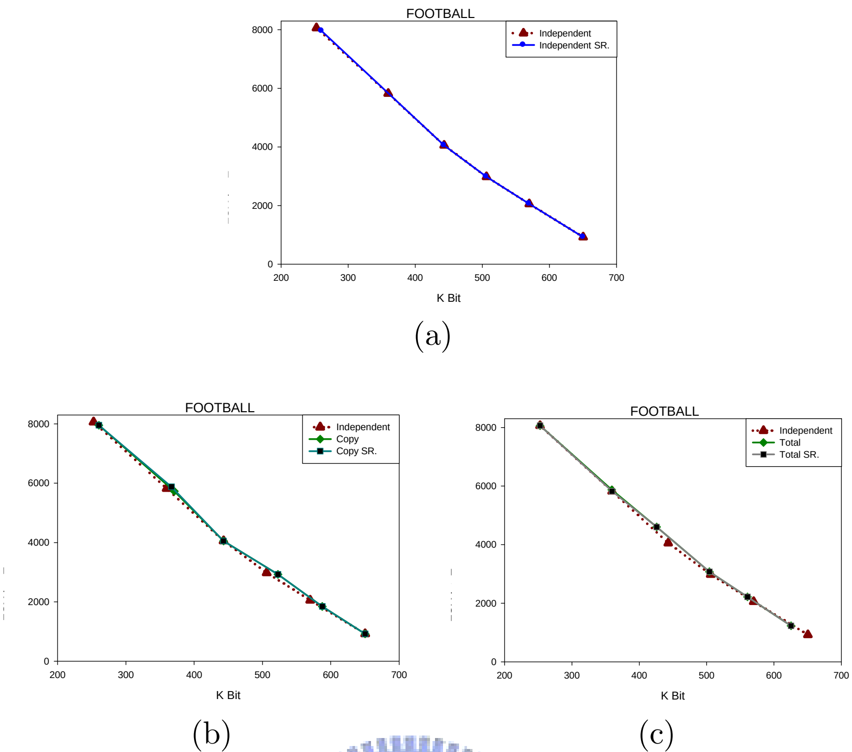


Figure 4.15: Comparison among proposed heuristic solution and its successive refinement(a), methodology 1 and its successive refinement(b), methodology 2 and its successive refinement(c) in Football.

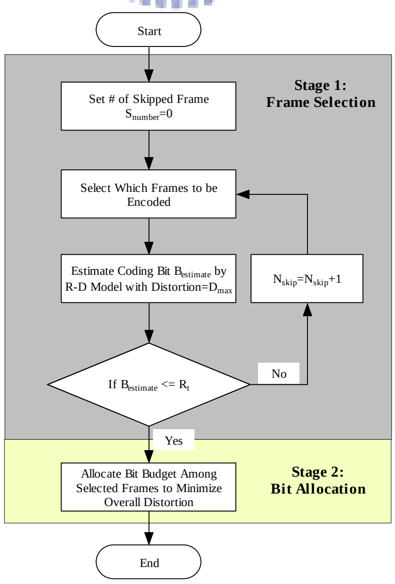


Figure 4.16: Proposed algorithm flowchart. At the first stage to decide which frames to be encoded; to allocate bit among selected frames at the second stage.

Step 3 :

Since skipped frames are out of the optimization target, we can simply allocate available bits among those selected coding frames and achieve the goal of minimum overall distortion.

CHAPTER 5

Conclusions and Future Works

In our work, we attempted to solve the quality- and rate-constrained problem and survey different constrained optimization problem methodologies. After comparing with related works and analyze our problem, we find out the best methodology for optimal solution is to use dynamic programming and the optimal path is a stairway-like curve, with huge complexity, the following lists our discoveries:

1. Constrained problems can be solved by dynamic programming. A simple algorithm for constrained optimization problem is to find the optimal solution among possible data set as integer programming; another commonly used mathematical tool for optimization problems is Lagrangian method, which turns the original constrained problem into unconstrained problem form.
2. For optmization problems with multiple constraints, each constraint corresponding to a Lagrangian multiplier, and the optimal solution should be obtained by iteratively adjusting each Lagrangian multiplier according to KKT conditions. As a result, the solution by constrained optimization problem like integer programming requires lower complexity than Lagrangian optimization; however, as for the single equality constrained problem, the Viterbi algorithm with Lagrangian

cost is commonly used owing to only one Lagrange multiplier variable to be solve, and the modern rate control apply this algorithm to real time system based on independency assumption.

3. According to optimal solution by integer programming for our problem, the optimal path is a stairway-like curve, which means number of frames to be encoded and coding frame selection should be different to optimize our constrained problem while the available bit budget changes. Furthermore, as long as more frame skipped, the MSE gap between each rate distortion line of different coding frame combination is larger; also, the interlacing part among coding bit range of different coding frame combination is getting smaller. From the above observation, it is a good choice to select another frame to be skipped while the available bit budget decreases in less frame skipped case for complexity issue, but it should choose another coding frame selection instead of skipping another frame immediately while more frame skipped.
4. In order to reduce complexity and to implement the proposed architecture in a real-time system, we try to develop a heuristic algorithm based on independent assumption and successive refinement and devide the solution into two stages:
   2.1) to find which frames to be encoded under rate constraint and distortion constraints;
   2.2) to allocate available bit budget among these chosen frames to achieve the objective of minimize total distortion.
   The experimental result shows this heuristic solution can reduce complexity exponentially, though it does not achieve a good performance because the optimal path reveals that the encoder should choose different frames to encode when target rate changes.

Our work is still in its early stage, we plan to extend our investigation in several directions:

1. To survey other adaptive frame skipping methods and continue to research if there exits a better algorithm for our problem.
2. To replace the frame copy with temporal interpolation methodology and analyze the effect.
3. To develop R-D models based on heuristic algorithm to allocate bits and minimize

total distortion for real-time systems.

Bibliography

[1] I. Ahmad and J. Luo, "On Using Game Theory to Optimize the Rate Control in Video Coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 2, Febuary 2006.

[2] J. G. David Forney, "The Viterbi Algorithm," *Proceedings of The IEEE*, vol. 61, no. 3, March 1973.

[3] W. Kaplan, "Maxima and Minima with Applications: Practical Optimization and Duality," *Wiley-IEEE*, 1999.

[4] S. Liu and C.-C. J. Kuo, "Joint Temporal-Spatial Bit Allocation for Video Coding With Dependency," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 1, January 2005.

[5] A. Ortega, "Optimal bit allocation under multiple rate constraints," *IEEE Data Compression Conference (DCC)*, 1996.

[6] A. Ortega, K. Ramchandran, and M. Vetterli, "Optimal Trellis-Based Buffered Compression and Fast Approximations," *IEEE Transactions on Image Processing*, vol. 3, no. 1, January 1994.

[7] K. Ramchandran, A. Ortega, and M. Vetterli, "Bit Allocation for Dependent Quantization with Applications to Multiresolution and MPEG Video Coders," *IEEE Transactions on Image Processing*, vol. 3, no. 5, September 1994.

[8] E. C. Reed and J. S. Lim, "Optimal Multidimensional Bit-Rate Control for Video Communication," *IEEE Transactions on Image Processing*, vol. 11, no. 8, August 2002.

[9] G. M. Schuster, G. Melnikov, and A. K. Katsaggelos, "A Review of the Minimum Maximum Criterion for Optimal Bit Allocation Among Dependent Quantizers," *IEEE Transactions on Multimedia*, vol. 1, no. 1, March 1999.

[10] Y. Shoham and A. Gersho, "Efficient Bit Allocation for an Arbitrary Set of Quantizers," *IEEE Transactions on Acoustics. Speech. and Signal Processing*, vol. 36, no. 9, September 1988.

[11] H. Song and C.-C. J. Kuo, "Rate Control for Low-Bit-Rate Video via Variable-Encoding Frame Rates," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 4, April 2001.

[12] K. Wang and J. W. Woods, "MPEG Motion Picture Coding With Long-Term Constraint on Distortion Variation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 3, March 2008.

-