MULTIPLE-DESCRIPTION LATTICE VECTOR QUANTIZATION FOR IMAGE AND VIDEO CODING BASED ON COINCIDING SIMILAR SUBLATTICES OF A_n

Ehsan Akhtarkavan

UNIVERSITI SAINS MALAYSIA

2012

MULTIPLE-DESCRIPTION LATTICE VECTOR QUANTIZATION FOR IMAGE AND VIDEO CODING BASED ON COINCIDING SIMILAR SUBLATTICES OF $\mathbf{A_n}$

by

Ehsan Akhtarkavan

Thesis submitted in fulfillment of the requirements
for the degree of
Doctor of Philosophy

April 2012

ACKNOWLEDGEMENT

First and foremost, I respectfully acknowledge my parents and family for everything they have done for the betterment of my life. Thank you very much, Mom and Dad.

Secondly, I am deeply indebted to my supervisor Dr. Mohd Fadzli Bin Mohd Salleh for all his supervision, guidance and efforts in providing a suitable environment for me to do this research and helping me to overcome all the problems I had encountered throughout this study.

.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AVC	Advanced video coding
AIR	Adaptive intra refresh
ARQ	Automatic repeat-request
Bpp	Bits per pixel
bps	Bits per second
CSL	Coinciding similar sublattice
CDF	Cohen-Daubechies-Feauveau
DWT	Discrete wavelet transform
EZW	Embedded zero-tree wavelet
GPRS	General packet radio service
GSM	Global system for mobile communications
HH	High-high
HL	High-low
LBG	Linde Buzo Gray
LVQ	lattice vector quantization
LL	Low-low
LH	Low-high
MD	Multiple descriptions
MDCLVQ	Multiple description coinciding lattice vector quantization
MDLVQ	Multiple description lattice vector quantization
MDSQ	Multiple description scalar quantization
Poset	Partially ordered set
PDF	Probability density function
PMF	Probability mass function
PSNR	Peak signal to noise ratio
RAMBL	Redundancy allocation at macro block level
RP-MDC	Redundant picture multiple descriptions coding
SPIHT	Set partitioning in hierarchical trees
SSL	Similar sub lattices
VoIP	Voice over internet protocol
VQ	Vector quantization

LIST OF SYMBOLS

Λ	A lattice
Λ'	A sublattice
λ	Lattice point
Λ'	Sublattice point
$\mathbb{R}$	Real number set
$\mathbb{Z}$	Integer number set
$\triangleq$	Equal
σ	Sigma
$\equiv$	Congruency relation for integer numbers
$\equiv'$	Coinciding congruency relation
$\mathcal{H}$	Hamiltonian algebra
K	The golden quadratic field
G^t	Transposition of the matrix G

PENGKUANTUMAN VEKTOR KEKISI BERBILANG KETERANGAN UNTUK PENGKODAN IMEJ DAN VIDEO BERPANDUKAN KESEKENAAN SUBKEKISI-SUBKEKISI A_n SERUPA

ABSTRAK

Pada hari ini penggunaan perhubungan multimedia boleh didapati di mana-mana. Sistem digital komunikasi digital menguruskan perwakilan data sama ada untuk penyimpanan atau penghantaran. Saiz data digital merupakan satu faktor yang penting untuk penghantaran data yang berkesan dan kebingkasan ralat merupakan faktor penting untuk sistem transmisi. Oleh itu, algoritma-algoritma pengekodan yang lebih berkesan dari segi mampatan dan kebingkasan ralat adalah amat diperlukan. Teknik ini mengelakkan penggunaan penghantaran semula data atau penghantaran semula automatik (ARQ) dalam sesuatu sistem rangkaian perhubungan. Pengekodan berbilang keterangan telah menjadi pilihan yang popular untuk penghantaran data yang lasak melalui saluran-saluran rangkaian tanpa kebolehharapan. Pengkuantuman vektor kekisi (LVQ) menghasilkan pengiraan yang lebih rendah untuk pemampatan data yang berkesan. Dalam tesis ini, pengkuantuman vektor kekisi berbilang keterangan untuk imej dan video berasaskan kesekenaan sub-kekisi serupa daripada A_n (MDCLVQ-A_n) telah disasarkan. Reka bentuk MDCLVQ adalah berdasarkan pada kesekenaan sub-kekisi serupa A_2 dan A_4. Kesekenaan sub-kekisi adalah sub-kekisi yang bergeometri serupa dengan indeks yang sama, tetapi dihasilkan oleh penjana matrik yang berlainan. Satu algoritma penandaan baru berdasarkan kesekenaan sub-kekisi juga telah dibangunkan. Skema MD yang

dicadangkan, MDCLVQ-A_2 dan MDCLVQ-A_4 telah diaplikasikan kepada pengekodan imej. Di samping itu, di dalam penyelidikan ini MDCLVQ-A_2 telah diaplikasikan kepada pengkodan video piawai H.264/AVC dan Motion JPEG2000 untuk membentuk skema-skema pengekodan video MD, MDCLVQ-H.264/AVC dan MDCLVQ-Motion JPEG2000. Skema pengekodan MD digunakan untuk meningkatkan keteguhan penghantaran melalui saluran-saluran cenderung ralat. Skema-skema MD yang dicadangkan telah diuji menggunakan dua jujukan imej dan lima video ujian piawai. Keputusan-keputusan ujikaji daripada aplikasi pengekod MD menunjukkan penambahbaikan dari segi prestasi pengekodan dan keteguhan penghantaran.

MULTIPLE-DESCRIPTION LATTICE VECTOR QUANTIZATION FOR IMAGE AND VIDEO CODING BASED ON COINCIDING SIMILAR SUBLATTICES OF A_n

ABSTRACT

Nowadays applications of multimedia communication are found everywhere. Digital communication systems deal with representation of digital data for either storage or transmission. The size of the digital data is a crucial factor for storage and error resiliency of the data is a crucial factor for transmission systems. Thus, it is required to have more efficient encoding algorithms in terms of compression and error resiliency. Multiple-description (MD) coding has been a popular choice for robust data transmission over unreliable network channels. This technique avoids having data retransmission or the automatic repeat-request (ARQ) in a communication system network. Lattice vector quantization (LVQ) provides lower computation for efficient data compression. In this thesis multiple-description lattice vector quantization for image and video coding based on coinciding similar sublattices of A_n (MDCLVQ-A_n) has been targeted. The design of the MDCLVQ is based on the coinciding similar sublattices of A_2 and A_4. The coinciding sublattices are geometrically similar sublattices with the same index, but generated by different generator matrices. A novel labeling algorithm based on the coinciding sublattices is also developed. The proposed MD coding schemes, MDCLVQ-A_2 and MDCLVQ-A_4 are applied to image coding. In addition, in this research the MDCLVQ-A_2 has been employed to H.264/AVC and Motion JPEG2000 video coding standards to

form MD video coding schemes, MDCLVQ-H.264/AVC and MDCLVQ-Motion JPEG2000. The MD coding schemes are used in order to increase the robustness of transmission over error-prone communication channels. The proposed MD coding schemes have been applied to two standard test images and five videos. The experimental results of application of the MD coders show improvements in terms of encoding performance and transmission robustness.

CHAPTER 1

INTRODUCTION

1.1 Preface

During the past forty years, image and video encoding for transmission has been focused by many researchers. At the beginning days, researchers' focus was on the development of analog methods for reducing image/video transmission bandwidth or bandwidth compression. After development of powerful digital computers and cheap integrated circuits, interests are shifted to digital compression approaches.

In recent years, according to widespread use of smart-phones and tablet personal computers, the demand for multimedia communication has been increased. Internet services such as the Internet telephony systems, Voice over Internet Protocol (VoIP), and audio/video streaming (e.g. online Radio/TV broadcasting) are becoming more and more popular. Clearly, many consumers enjoy the Internet telephony services provided for free by many different companies such as Google Talk™, Yahoo Messenger™, Oovoo™, and Skype™. This trend is steadily growing, and more and more people are switching their landline phones into VoIP compatible

telephones. On the wireless side, it is likely that cell phones soon are to use the same Internet telephony systems using the General Packet Radio Service (GPRS) (Keshav, 2005) which is available on 2G, 3G, and 4G cellular communication systems of the global system for mobile communications (GSM).

These types of "real-time" services often require low delay, high error resiliency, low packet-loss rates, and high bandwidth in order to deliver the quality desired by the end users. However, current packet-switched networks and communication infrastructures do not guarantee such needs and therefore the desired quality of service may not be achieved. Thus, this reveals the need for encoding schemes that offer high compression ratios and robust transmissions over unreliable wireless and wired network channels.

Digital communication channels and digital storage systems usually suffer from limited available bandwidth and limited storage capacity. Thus, digital compression is an essential part of any information processing systems. The term data compression refers to the process in which the amount of data required to represent a given quantity of information is reduced. Data is the physical material that is used to contain and represent given information, therefore several different data may represent the same information.

The amount of data that is required to represent the information varies from one representation to another and there may be methods to convert one representation to another representation with fewer amounts of data. Since these data representations are equal in terms of the information they represent, the amount of data that is omitted in the conversion (compression) process is called redundancy.

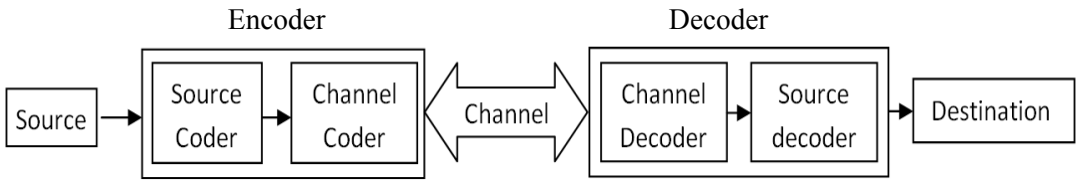

Figure 1-1: The encoder-channel-decoder model and the sub-blocks.

A conversion system may contain two building blocks: an encoder and a decoder. The encoder is responsible for converting from the representation with higher data to the representation with less data and the decoder does the reverse conversion. The encoder-decoder couple is usually connected through a channel, that is, encoder-channel-decoder. Therefore, each of these two blocks contains two different sub-blocks with independent behaviors. The encoder is made up of a source encoder and a channel encoder. The source encoder is responsible for data compression and the channel coder is responsible for error resiliency. The decoder is made up of a source decoder and a channel decoder. In fact, the conversion and the compression happen in both sub-blocks. The encoder-channel-decoder model and the sub-blocks are depicted in Figure 1-1. Multiple-descriptions lattice vector quantization is a method to address both error resiliency and data compression.

1.2 Problem statement

The demand for higher data communication rates has risen and this introduces a problem because the available bandwidth becomes scarce. In addition, current networks are error prone. Thus, there is a great demand for efficient encoding techniques that offer error resiliency and data compression. In single-channel communications, channel impairment reduces the communication performance. Therefore, multiple-channel data transmission techniques are much more demanded for reliable data communications.

The packet switching networks enable the user to increase the robustness in terms of delay and packet loss by exploiting diversity. For example, every packet may be duplicated and transmitted over two different paths, throughout the network. If one of the packets is lost, there will be no reduction in quality at the receiving side. Thus, there is a great degree of robustness. On the other hand, if none of the channels fail, no packets will be lost. Thus, there is no point in using both packets. Besides, the robustness that roots from diversity poses an extra cost.

However, if small quality degradation is compromised with a lower transmission cost, when receiving one description, and preserving a good quality on reception of both packets, it is possible to reach robustness without extra cost. This idea of trading off required bit rate vs. quality between a numbers of packets (descriptions) is usually referred to as the Multiple-Description (MD) coding. A typical 2-channel MD coding scheme includes two joint encoders, one central decoder and two side decoders. In such scheme, a source that requires R b/s for encoding is encoded by the two joint encoders that require R_1 and R_2, provided that $R \geq R_1+R_2$. The joint encoders generate two descriptions of the source so that the central decoder can provide high fidelity if both descriptions are received and side decoders can provide an acceptable approximation of the source if only one description is received.

In single description quantization, an extra transmission bit rate reduces the squared error distortion by a factor of 4 (Vaishampayan et al., 2001). However in multiple-description quantization if the transmission rate is increased by 0.5 bit then the central distortion and side distortions are decreased by $2^{-(1+a)}$ and $2^{-(1-a)}$, respectively, for any $a \in (0,1)$. This means that by adjusting the value of a, the

central distortion and the side distortions are decreased by different amounts but the product of the distortions decreases by a factor of 4 (Vaishampayan et al., 2001).

In addition, the vector quantization decreases the granular distortions because in higher dimensions it is possible to construct more spherical Voronoi cells than the hypercube (Vaishampayan et al., 2001). However, the vector quantization technique usually comes with high design complexity and computational load such as Linde Buzo Gray (LBG) algorithm (Linde et al., 1980). Therefore, the second problem is to have a quantizer with lower complexity, such as lattice vector quantization.

The MD coding using lattice codebooks or multiple-description lattice vector quantization (MDLVQ) has appeared to be an attractive scheme to tackle network failures and increase robustness of the multimedia communications especially for those applications that retransmission is not possible or is costly (Vaishampayan et al., 2001). The proposed scheme in Vaishampayan et al. (2001), SVS-MDLVQ, uses a labeling function to map lattice points to two point of a sublattice using a labeling function. The optimized-MDLVQ (Bai et al., 2007) proposes an optimized labeling function for SVS-MDLVQ. But the descriptions generated by the optimized labeling function are not balanced and even alternative transmission over the two channels cannot impose balance between the descriptions. Thus, it is required to use inherent symmetries of the A_n lattices to develop new labeling functions that generate balanced descriptions

1.3 Objectives of the research

This research focuses on developing new MD coding schemes using the lattice codebook to address channel failure without re-transmission. The key objectives of this work are given as follows:

1. To investigate inherent symmetries and properties of A_2 and A_4 lattices.
2. To develop the mathematics for constructing the coinciding similar sublattices of A_2 and A_4.
3. To develop multiple-description lattice vector quantization schemes, based on the coinciding similar sublattices of A_2 and A_4, respectively.
4. To apply the proposed MD coding schemes to standard test images and videos and compare the results with the state-of-the-art MD coding schemes.

1.4 The scope of the research

Diversity systems are potential methods that can be used to solve the problem of retransmission, automatic repeat-request (ARQ) in communication networks. Besides, it adds error resiliency to communication systems by sending information over multiple channels. This is due to their probabilities of failure on every channel are independent events. Thus, the probability of receiving at least one of the channels is greatly increased. The MD coding is an efficient data transmission method using for transmission using multiple communication channels. In such a system, the MD coding is designed to encode the baseband source by two joint encoders into two bitstreams. If both channels are received, a high-quality reconstruction can be achieved. However, if one channel is corrupted, an acceptable degraded reconstruction can be achieved from the other channel. The MD coding eliminates the need for data retransmission or ARQ and offers communication networks with higher error resiliency.

This study focuses on the design and development of a new data transmission system for application in communication networks based on a particular MD coding technique that incorporates lattice vector quantization. Several factors are important

in multiple description lattice vector quantization (MDLVQ) schemes such as the required bit rate for encoding the two bitstreams. The required bit rate is affected by the value of fundamental area of the lattice.

The second factor is the trade-off between the reconstruction qualities in the central decoder versus the reconstruction qualities in the side decoders. The performances of the MD coding schemes are tested using the baseband data. In this case, the simplest and most available baseband data for this particular application are image and video data as commonly found in the related literatures.

1.5 Thesis outline

An introduction to requirements of data transmission systems, error resiliency and compression efficiency were presented in Section 1.1. In addition, the application of MD coding to address these requirements were presented. In Section 1.2, the definition of the MD coding was presented and the problem statement was described. The objectives of the research and the scope of the research were presented in Section 1.3 and Section 1.4, respectively.

In Chapter 2, preliminary concepts related to the MD coding and MDLVQ are presented. In addition, brief reviews of the MD coding literature as well as current state-of-the-art MD coding schemes are presented. Chapter 3 includes the details of the development of the coinciding similar sublattices of A_2 and A_4. In this chapter, the required mathematics for calculating the generator matrices of the coinciding similar sublattices of A_2 and A_4 and their corresponding transformation matrices are provided. In addition, a new labeling function based on the coinciding similar sublattices of A_2 is proposed for MDCLVQ scheme.

Chapter 4 is devoted to the design and development of the multiple-description coinciding lattice vector quantization (MDCLVQ) schemes for MD image coding, MDCLVQ-A_2 and MDCLVQ-A_4. In addition, in this chapter implementation of the proposed multiple-description video coding schemes, MDCLVQ-H.264/AVC and MDCLVQ-Motion JPEG2000 are presented.

Experimental results related to the application of the proposed MDCLVQ schemes based on the coinciding similar sublattices of A_2 and A_4 to standard test images and standard test video sequences are provided in Chapter 5. In addition, the performances of the proposed schemes in terms of the reconstruction quality and bit rate efficiency are provided in Chapter 5. In Chapter 6, the thesis is concluded and guidelines for future researches are discussed.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Digital signal processing for various applications such as video/image communication, wireless communication, and biomedical technologies has been found interesting by many researchers. Increasing number of wireless communication users has increased the competition in using the transmission bandwidth and hence tightened the constraints in bandwidth allocation. In addition, the error prone networks require error resilient coding schemes. Therefore, in order to reduce the amount of data to be transmitted, compression techniques and error resilient techniques are very much needed. Multiple-descriptions lattice vector quantization is a method that offers data compression and error resiliency.

In this chapter, different techniques that have been developed in order to attack the mentioned challenges such as MD coding, lattice vector quantization (LVQ) and multiple-description lattice vector quantization (MDLVQ) are presented. However, there are several preliminary concepts that need to be discussed first. In Section 2.2.1 the elementary definitions and properties of the lattices are presented. Then, the

geometrically similar sublattices are described in Section 2.2.2. In Section 2.2.3, the vector quantization is presented and lattice vector quantization is described in Section 2.2.4. Fast quantizing algorithms are common lattice quantizing algorithms and they are presented in Section 2.2.5. The MD coding technique is described in 2.2.6. Rate and distortion theory for MDLVQ are presented in Section 2.2.7. In Section 2.2.8, the discrete wavelet transform is discussed and the details of the quaternion algebra are presented in Section 2.2.9. The entropy coding is presented in Section 2.2.10. These serve as the background knowledge for the reader. These techniques will be used as the main topics in the rest of this thesis. The literature review of different types of techniques related to MD coding using scalar quantization and the important MDLVQ schemes proposed so far are presented in Section 2.3.1 and 2.3.2, respectively.

2.2 Background

2.2.1 Elementary of lattices

In mathematics (algebra), a lattice is defined as a partially ordered set (poset) in which any two elements have a unique supremum (the element's least upper bound or join) and an infimum (greatest lower bound or meet). In other words, a lattice is considered as a subset of points in the Euclidean space that share a common property. For example the lattice A_n is a subset of points with n+1 coordinates, such that the sum of these coordinates is zero. Therefore, the lattice A_n can be defined as:

$$A_n = \{(x_0, x_1, \dots, x_n) \in Z^{n+1} : x_0 + x_1 + \cdots + x_n = 0\} \qquad (2\text{-}1)$$

An n-dimensional lattice Λ in $\mathbb{R}^n$ is denoted by $\Lambda = \langle \boldsymbol{b_1}, \boldsymbol{b_2}, \dots, \boldsymbol{b_n} \rangle$. It means that Λ consists of all integer linear combinations of a basis vectors $\{\boldsymbol{b_1}, \boldsymbol{b_2}, \dots, \boldsymbol{b_n}\}$ in

$\mathbb{R}^n$ (Heuer, 2008). Thus, the fundamental parallelotope of a lattice Λ is defined as (Conway and Sloane, 1998)

$$\theta_1 \boldsymbol{b_1} + \theta_2 \boldsymbol{b_2} + \cdots + \theta_n \boldsymbol{b_n} \; (0 \leq \theta_i < 1) \tag{2-2}$$

The fundamental parallelotope is the building block of the lattice because if it is repeated many times, the whole space is filled in a way that there is only one lattice point in each parallelotope. There are many ways of choosing a basis and a fundamental parallelotope for a lattice Λ. But the volume of the fundamental region is uniquely determined by Λ, and the square of this volume is called the determinant of the lattice (Conway and Sloane, 1998).

The lattice points are generated using a generator matrix. The generator matrix is composed of the basis vectors of the lattice. The generator matrix of the lattice Λ with the basis vectors $\boldsymbol{b_1} = (b_{11}, b_{12}, \dots, b_{1m})$, $\boldsymbol{b_2} = (b_{21}, b_{22}, \dots, b_{2m})$, $\dots, \boldsymbol{b_n} = (b_{n1}, b_{n2}, \dots, b_{nm})$ is given as (Conway and Sloane, 1998):

$$\boldsymbol{G} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{pmatrix} \tag{2-3}$$

The Grammian or Gramm matrix of a lattice Λ is defined as $\boldsymbol{A} = \boldsymbol{G}\boldsymbol{G^t}$, where $\boldsymbol{G^t}$ is the transposed matrix of $\boldsymbol{G}$. The Gramm matrix determines the linear independence of the basis vectors, that is, they are linearly independent if and only if the determinant of the Gram matrix is non-zero. Two lattices are called equivalent if they have the same Gramm matrix or if the Gramm matrices are proportionate. The determinant of a lattice Λ is also equal to the determinant of the Gramm matrix (Conway and Sloane, 1998):

$$det\,\Lambda = det\,\boldsymbol{A} \tag{2-4}$$

If the generator matrix is a square matrix then Eq. (2-4) is written

$$det\,\Lambda = (det\,\boldsymbol{G})^2 \tag{2-5}$$

Thus, the volume of the fundamental parallelotope of a lattice Λ is also calculated as (Conway and Sloane, 1998)

$$vol = \det \boldsymbol{G} = \sqrt{\det \Lambda} = \sqrt{\det \boldsymbol{A}} \tag{2-6}$$

For example, the hexagonal lattice is a subset of the complex space $\mathbb{C}$, and at unit scale it is generated by the basis vectors $\{\mathbf{1}, \boldsymbol{\omega}\} \subset \mathbb{C}$, where $\boldsymbol{\omega} = -1/2 + i\sqrt{3}/2$ (Vaishampayan et al., 2001). Therefore, the hexagonal lattice at unit scale is generated by

$$\boldsymbol{G}_{2\times 2} = \begin{pmatrix} Re(\mathbf{1}) & Im(\mathbf{1}) \\ Re(\boldsymbol{\omega}) & Im(\boldsymbol{\omega}) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{-1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \tag{2-7}$$

and the Gramm matrix of $\boldsymbol{G}_{\mathbf{2\times 2}}$ is calculated as (Conway and Sloane, 1998)

$$\mathbf{A_{2x2}} = \boldsymbol{G}_{2\times 2}\boldsymbol{G}_{2\times 2}{}^{\mathbf{t}} = \begin{pmatrix} 1 & \frac{-1}{2} \\ \frac{-1}{2} & 1 \end{pmatrix} \tag{2-8}$$

and $det\,\Lambda_{2\times 2} = det\,\boldsymbol{A}_{\mathbf{2\times 2}} = \frac{3}{4}$. Thus, the volume (or area) of fundamental parallelotope of Λ will be calculated as $vol = \sqrt{det\,\Lambda_{2\times 2}} = \frac{\sqrt{3}}{2}$. The hexagonal lattice generated by Eq. (2-7) and its fundamental parallelotope are shown in Figure 2-1.

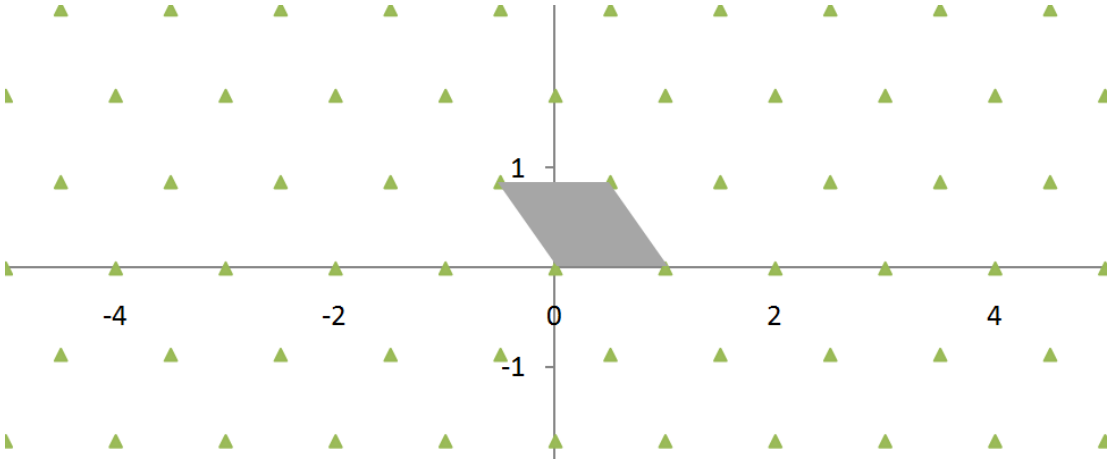

Figure 2-1: The hexagonal lattice and its fundamental parallelotope.

The hexagonal lattice is also generated by

$$\boldsymbol{G}_{2\times 3} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \tag{2-9}$$

and the Gramm matrix of $\boldsymbol{G}_{2\times 3}$ can be calculated as

$$\boldsymbol{A}_{2\times 3} = \boldsymbol{G}_{2\times 3}\boldsymbol{G}_{2\times 3}{}^{\mathrm{t}} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = 2\mathbf{A}_{2\times 2} \tag{2-10}$$

According to $\boldsymbol{A}_{2\times 3} = 2\boldsymbol{A}_{2\times 2}$, the lattices generated by $\boldsymbol{G}_{2\times 2}$ and $\boldsymbol{G}_{2\times 3}$ are equivalent lattices. However, the determinant of $\Lambda_{2\times 3}$ is 3 and the volume of fundamental parallelotope of $\Lambda_{2\times 3}$ is $\sqrt{3}$. This is because $\boldsymbol{G}_{2\times 2}$ and $\boldsymbol{G}_{2\times 3}$ both describe the hexagonal lattice but in different coordinates and on different scales (Conway and Sloane, 1998).

In an n-dimensional lattice Λ, the Voronoi region of a lattice point is defined as the union of all non-lattice points within $\mathbb{R}^n$ that are closer to this particular lattice point than any other lattice point. Thus, the Voronoi region of $\lambda \in \Lambda$ is defined as (Vaishampayan et al., 2001)

$$V(\lambda) \triangleq \{x \in \mathbb{R}^{\mathrm{n}} : \|x - \lambda\| \leq \|x - \lambda'\|, \forall \lambda' \in \Lambda\} \tag{2-11}$$

As a consequence, all the points within $V(\lambda)$ must be quantized to λ. The Voronoi regions of the points in the A_2 are hexagons; therefore, it is called the hexagonal lattice. The Voronoi region of a sublattice point λ' is the set of all lattice points that are closer to λ' than any other sublattice points. Thus, the Voronoi region of $\lambda' \in \Lambda'$ is defined as

$$V(\lambda') \triangleq \left\{\lambda \in \Lambda : \|\lambda - \lambda'\| \leq \left\|\lambda - \lambda''\right\|, \forall \lambda'' \in \Lambda'\right\} \tag{2-12}$$

In order to define an n-dimensional lattice Λ, it is enough to specify its fundamental area. However, it is usually expressed through its dimensionless second moment of inertia. The dimensionless second moment of inertia $DSMI(\Lambda)$ is defined as (Conway and Sloane, 1998)

$$DSMI(\Lambda) \triangleq \frac{1}{vol^{1+\frac{2}{n}}} \int_{V(0)} \|x\|^2 \, dx \tag{2-13}$$

where $\|x\|^2 = \frac{1}{n}\sum_{i=1}^{n} x_i . x_i$ is the l_2-norm of x, vol is the fundamental volume of Λ, and $V(0)$ is the Voronoi region of the origin (Conway and Sloane, 1998).

2.2.2 Geometrically similar sublattices of A_2

Assume that Λ is an L-dimensional lattice with the generator matrix $\boldsymbol{G}$. A sublattice $\Lambda' \subset \Lambda$ with generator matrix $\boldsymbol{G}'$ is said to be geometrically similar to Λ if and only if $\boldsymbol{G}' = c\boldsymbol{UGB}$, for nonzero scalar c, an integer matrix $\boldsymbol{U}$ with $\det \boldsymbol{U} = \pm 1$, and a real orthogonal matrix $\boldsymbol{B}$ (with $\boldsymbol{BB^t} = \boldsymbol{I}$) (Conway and Sloane, 1998).

The index N is defined as the ratio of the fundamental volume of the sublattice Λ' to the fundamental volume of the lattice Λ. The fundamental volume of the lattice (vol) is equal to the determinant of the generator. Thus, N is calculated by

$$N = \frac{vol'}{vol} = \sqrt{\frac{det\,\Lambda'}{det\,\Lambda}} = \frac{det\,\boldsymbol{G'}}{det\,\boldsymbol{G}} \tag{2-14}$$

In other words, N is the number of lattice points within the Voronoi region of the sublattice points. Therefore, the value of N controls the coarse degree of the sublattice as well as the amount of redundancy in the MD coder (Vaishampayan et al., 2001). It has been shown in Bernstein et al. (1997) and Vaishampayan et al. (2001) that, for the hexagonal lattice, Λ' is similar to Λ if N is of the form

$$N = \alpha^2 - \alpha\beta + \beta^2 \text{ for } \alpha, \beta \in \mathbb{Z} \tag{2-15}$$

In addition, N must be in the form of $N = \sum_{i=0}^{K} n_i$, where, n_i denotes the number of points at squared distance i from the origin. If these conditions are met then the basis vector of the sublattice Λ' will be $\boldsymbol{u} = \alpha + \beta\boldsymbol{\omega}$ and $\boldsymbol{v} = (\alpha + \beta\boldsymbol{\omega})\boldsymbol{\omega}$. Sublattice $\Lambda' \subset \Lambda$ is considered as a clean sublattice if all the points of the Λ reside only inside the Voronoi region of the sublattice points rather than on the boundary of the Voronoi region (Conway et al., 1999). It means that the lattice points are not shared between the Voronoi regions of adjacent sublattice points.

The sublattices of A_2 are clean, if and only if, α and β are relatively primes. It follows that A_2 has a clean similar sublattice of index N if and only if N is a product of primes congruent to 1 (mod 6) (Conway et al., 1999). The sequence of integers that generate clean sublattices of the hexagonal lattice are named A038590 by Sloane (2000). In other words, α and β are selected such that the value of N satisfies these

conditions and hence a clean similar sublattice of the hexagonal is generated. For example, with $\alpha = -3$ and $\beta = 2$, a clean similar sublattice of the hexagonal lattice with index $N = (-3)^2 - (-3)(2) + (2)^2 = 19$ is generated. The basis vectors are calculated as $\boldsymbol{u} = (-3) + (2)\boldsymbol{\omega} = -4 + i\sqrt{3}$ and $\boldsymbol{v} = \left(-4 + i\sqrt{3}\right)\boldsymbol{\omega} = 0.5 - 2.5i\sqrt{3}$. Thus, the corresponding generator matrix will be calculated as

$$\boldsymbol{G}' = \begin{pmatrix} Re(\boldsymbol{u}) & Im(\boldsymbol{u}) \\ Re(\boldsymbol{v}) & Im(\boldsymbol{v}) \end{pmatrix} = \begin{pmatrix} -4 & \sqrt{3} \\ 0.5 & -2.5\sqrt{3} \end{pmatrix} \tag{2-16}$$

It is also possible to calculate the index of the sublattice generated by $\boldsymbol{G}'$ using Eq. (2-14). The determinant of the generator of the hexagonal lattice at unit scale is $\sqrt{3}/2$. The determinant of $\boldsymbol{G}'$ is calculated as $det(\boldsymbol{G}') = (-4) \times \left(-2.5\sqrt{3}\right) - (0.5) \times (\sqrt{3}) = 19\sqrt{3}/2$. Thus, the index will be $(19\sqrt{3}/2)/(\sqrt{3}/2) = 19$. The sublattice generated by $\boldsymbol{G}'$ is shown in Figure 2-2 with blue squares and the hexagonal lattice points are shown with light blue triangles. The fundamental parallelotope of the hexagonal lattice and the similar sublattice generated by $\boldsymbol{G}'$ are shown with small and big parallelogram, respectively.

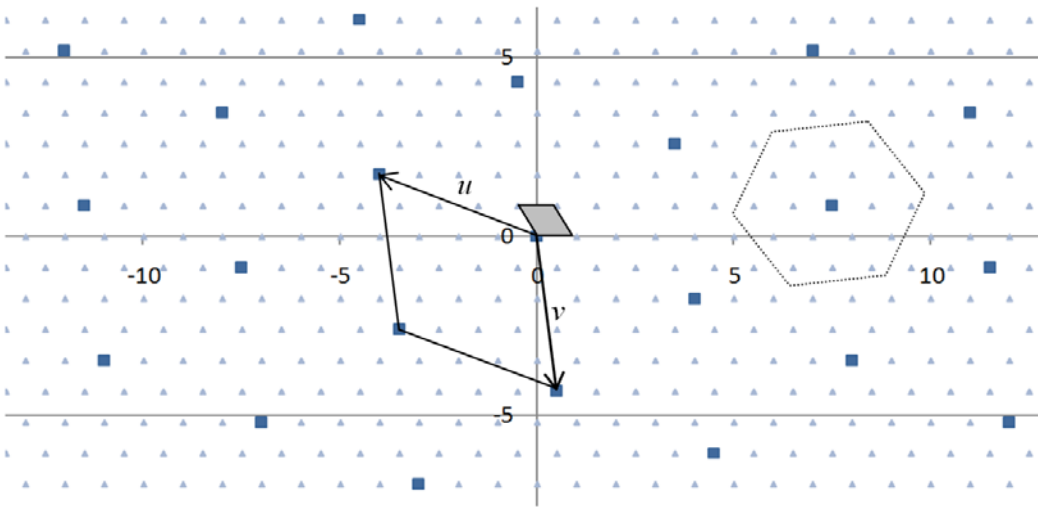

Figure 2-2: The geometrically similar sublattice of A_2 with index $N=19$ generated by Eq. (2-16).

The basis vectors, $\boldsymbol{u}$ and $\boldsymbol{v}$, are also shown. The Voronoi region of the sublattice point $\lambda' = (7.5, 0.5\sqrt{3})$ is shown with a dashed hexagon. It is seen in Figure 2-2 that $\boldsymbol{G'}$ has generated a clean sublattice because there are no lattice points on the boundary of the Voronoi region of the sublattice points.

2.2.3 Vector quantization

Quantization algorithms fall into two different categories namely: scalar quantization and vector quantization. Each of them has its own models. Scalar quantization is very similar to the rounding up algorithm used in analog to digital transformers. Instead of representing the data with uncountable continuous amounts, it simply creates a countable set of function discontinuities called the decision and reconstruction levels of the quantizer. In a very simple approach scalar quantization is nothing but a division and dequantization is nothing but a multiplication. In other words quantization is considered as:

$$Quantized\_number = \frac{Original\_number}{Quantization\_factor} \tag{2-17}$$

However, because the division usually comes with a loss of precision the inverse quantization or the dequantization may not produce the original number but generate an approximation of the original number:

$$Original\_number \cong Quantized\_number \times Quantization\_factor \tag{2-18}$$

In vector quantization the input stream of data are vectorized and then the vectors are quantized. For example for two dimensions, every two input data are considered as a vector and then mapped into a two dimensional plane like the Cartesian coordinate system. Therefore, the quantization error in vector quantization

is higher than the scalar quantization. Vector Quantization (VQ) technique has been employed in applications related to multimedia communications (Gersho and Gray, 1992). The VQ maps a group of input data to a fixed codeword which is readily available in a codebook.

In the early works related to VQ, many researchers set their attention on VQ techniques where the codebooks are generated using algorithms such as "an algorithm to compute the nearest point in the lattice A_n^*" in McKilliam et al. (2008) and "finding the closest lattice point by iterative slicing" in Sommer et al. (2007). The VQ technique is explored in order to have higher compression ratio. The work in Patrick and Christine (2001) presents vector indexing algorithms to generate the codebook that enables one to trade the codebook size for arithmetic operations. In Linde et al. (1980) an algorithm for generating an accurate codebook has been proposed. However, the codebook generation requires huge amount of computation. Thus, VQ techniques such as Lattice Vector Quantization (LVQ) that offer low complexities and good quantization performances are used (Chen, 2008, Servetto et al., 1998, Bai et al., 2007, Vasuki and Vanathi, 2006).

2.2.4 Lattice vector quantization

Lattice vector quantization (LVQ) is a vector quantization technique that reduces the amount of computation for codebook generation since the lattices have regular structures. A finite set of points $y_1, \dots, y_M$ in an n-dimensional Euclidean space, R^n , is called an Euclidean code (Conway and Sloane, 1982b). An n-dimensional quantizer is a mapping function $\mathrm{Q} : R^n \longrightarrow R^n$ that sends each point $x \in R^n$ into $Q(x)$ provided that $Q(x)$ is the nearest code point. The code points may be

selected according to any type of relationship. If the code points are selected from a lattice, then the quantizer would be called a lattice vector quantizer.

2.2.5 Fast quantizing algorithms for A_2

Fast quantizing algorithms are a family of lattice vector quantization algorithms presented in Conway and Sloane (1982a) for different root lattices. The quantization using A_n lattice points is a projection from n-dimensional space onto $\sum_{i=1}^{n+1} x_i = 0, x_i \in \mathbb{Z}$ hyper plane. The fast quantizing algorithm first projects the n-dimensional input vector onto n+1 dimensional vectors on $\sum_{i=1}^{n+1} x_i = 0, x_i \in \mathbb{R}$ hyper plane using a matrix-multiplication Conway and Sloane (1982a). Then, using a manipulation the projected point is mapped onto a lattice point.

For example consider quantization using A_2 lattice points. The input stream of data is vectorized into 2-dimensional vectors. Then, each input vector (i_1, i_2) is projected onto the 3-dimensional space, $(x_0, x_1, x_2) \in \mathbb{Z}^3$ with constraint that $x_0 + x_1 + x_2 = 0$. In order to do the projection (i_1, i_2) onto the 3-dimensional space, it is multiplied on the right by the transformation matrix **T** given as (Conway and Sloane, 1982a)

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & -1 \\ 1/\sqrt{3} & -2/\sqrt{3} & 1/\sqrt{3} \end{pmatrix} \tag{2-19}$$

If the expression $x_0 + x_1 + x_2 = 0$ does not hold, all the coordinates need to be rounded to the nearest integer points, while keeping the original values in another variable. The projected 3-dimensional vector is easily quantized (mapped) to the nearest lattice point by a simple manipulation. The sum of the differences between each coordinate of the original projected point to the nearest integer is calculated. If

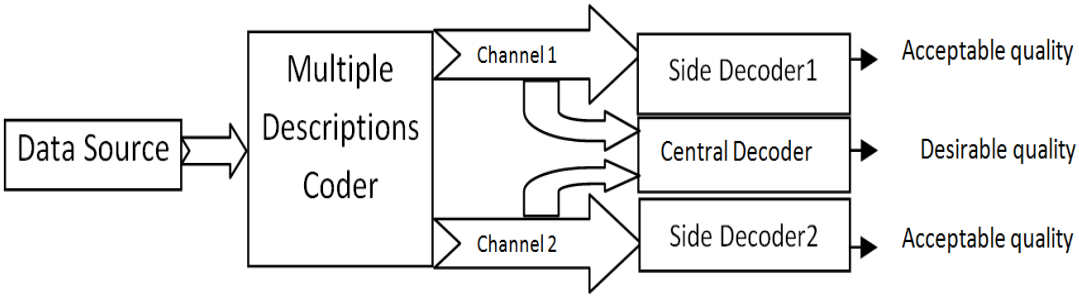

Figure 2-3: A general scheme of the MD coding scheme.

the sum of the differences is positive, then 1 is subtracted from the coordinate farthest from the integer. On the other hand, if the sum is negative, then 1 is added to the coordinate with the most difference. Thus, performing the computation-intensive nearest neighboring search algorithm is avoided. The two-dimensional version of the result point is calculated by right multiplying (x_1', x_2', x_3') by $\frac{1}{2}\boldsymbol{T^t}$ (Conway and Sloane, 1982a). The development of the lattice vector quantizer will be presented in Chapter 3.

2.2.6 Multiple-description coding

Multiple-description (MD) coding is a method to address network impairments when the re-transmission is expensive or impossible. According to Goyal (2001), MD coding can effectively address packet loss without the need for retransmission, thus this meets the network requirements. In this scheme, a stream of input data is transformed into several different independent descriptions and sent over different channels of a diversity system. At the receiver if all the descriptions are received correctly, the original data will be reconstructed accurately. But, in case some of the descriptions fail to reach the destination, due to channel failure, the rest of the descriptions, which are fed via side decoders, are used to find an estimate of the original data. The performance of the MD system to reconstruct the original data can

be of several levels of accuracy. A typical 2-channel MD coding scheme is shown in Figure 2-3. In this scheme, a source that requires R b/s for encoding is encoded by two joint encoders that require R_1 and R_2, provided that $R \geq R_1+R_2$.

For example, consider a simple two channels scheme in which both descriptions are the same. If either description is lost then the other would be useful. However, if both descriptions are available then one will be useless and hence the bandwidth has been wasted. In other word, receiving more descriptions must result in better reconstruction quality which can be offered by the MD coding (Goyal, 2001). According to information theoretic approach, the MD coding scheme may not require more bit rate (bandwidth) than the single description system. In the MD coding system, there is always a trade-off between the required bit rate and the distortion. Thus, in MD coding scheme, compression efficiency is sacrificed in order to gain error resiliency. Therefore, MD coding should be applied only if it does not require too extra bit rate or a wider bandwidth (Goyal, 2001).

2.2.7 Rate and distortion theory for MDLVQ

Information theory is the basis of the contemporary digital communication networks. Shannon first introduced concepts of information theory in his famous channel coding theorems in Shannon (1948). In order to describe an arbitrary real number, an infinite number of bits are required; therefore representing it with a finite number of bits, as in digital systems, can never be perfect. In rate distortion theory, the main goal is to determine the minimum expected distortion achievable at a particular rate for a given source (Cover and Thomas, 1991).

Assume the source X that generates independent and identically distributed random variables has a finite number of alphabets $X^n = \{X_1, X_2, \dots, X_n\}$. A mapping

that maps every source symbol into a finite set of code symbols is called a source encoder. These symbols are called source symbols. The number of symbols within the finite set of code symbols, denoted as n, impacts the required bit rate R for encoding the source. The encoder assigns an index $f_n(X^n) \in \{1,2,\dots,2^R\}$ to every code symbol. The decoder is responsible for reconstructing the original source symbol. However, the reconstructed values $\hat{X} = \{\hat{X}_1, \hat{X}_2, \dots, \hat{X}_n\}$ may not be the same as the original values (Cover and Thomas, 1991). A distortion function d is defined as a mapping from the set of Cartesian product of the source alphabet set and reproduction alphabet set into a non-negative real number

$$d: X \times \hat{X} \to \mathbb{R}^+ \tag{2-20}$$

The difference between X_i and $\hat{X}_i$ is the error introduced by the source encoder. The amount of error would validate the accuracy of the reconstruction. If the error becomes more than an agreed upon threshold then the data would be useless. Thus, it is important to find a measure to determine the amount of error.

The quality of reconstruction is usually measured by the distortion between the original data and the reconstructed data. Distortion is a measure that shows how well a source letter is reconstructed using its representative (Goyal, 2001). The distortion between a source letter X and its reconstructed version $\hat{X}$ is calculated as

$$dist\left(X^n, \hat{X}^n\right) = \frac{1}{n}\sum_{i=1}^{n} d\left(X_i, \hat{X}_i\right), \tag{2-21}$$

where d may be any distortion measure. The squared error is the most common distortion measure. The squared error is calculated as

$$d(x, \hat{x}) = (x - \hat{x})^2 \tag{2-22}$$

Thus, the general distortion between a typical source code and the reproduction is defined as the expected value of the $dist(x, \hat{x})$. This distortion is defined as

$$D = E\left[dist\left(X^n, \widehat{X^n}\right)\right] = E[dist(X^n, g(f(X^n)))] \tag{2-23}$$

where, $f(x)$ is the source coder and $g(x)$ is the source decoder.

In MD coding systems, there are two types of decoders and hence two types of distortions must be defined namely central decoder and side decoders. The central decoder is used whenever all the descriptions are received correctly and the side decoders are used whenever one or more descriptions are lost. For example, if the MD coding scheme includes two transmission channels, then the central decoder needs both descriptions to be received. However, the side decoders may decode using either first or second descriptions. The central distortion or the distortion related to the central decoder of a MDLVQ scheme is defined as

$$d_c \triangleq \sum_{\lambda_c \in \Lambda_c} \int_{V_\Lambda(\lambda_c)} \|x - \lambda_c\|^2 f_X(x)\, dx \approx DSMI(\Lambda_c)\, vol^{2/L} \tag{2-24}$$

where $V_\Lambda(\lambda)$ is the Voronoi region of λ, vol is the volume of fundamental parallelotope of the L-dimensional lattice, $DSMI(\Lambda_c)$ is the dimensionless second moment of inertia of the central lattice Λ_c Eq. (2-13), and f_X is the L-fold probability distribution function given by (Ostergaard et al., 2006)

$$f_X(x) = \prod_{i=0}^{L-1} f(x_i) \tag{2-25}$$

The distortion of the ith side decoder is defined as (Ostergaard et al., 2006)

$$d_i = \sum_{\lambda_c \in \Lambda_c} \int_{V_\Lambda(\lambda_c)} \|x - a_i(\lambda_c)\|^2 f_X(x)\, dx$$

$$\approx d_c + \sum_{\lambda_c \in \Lambda_c} \int_{V_\Lambda(\lambda_c)} \|\lambda_c - a_i(\lambda_c)\|^2 P(\lambda_c)\, dx \qquad (2\text{-}26)$$

where $\alpha_i(\lambda)$ is the component mapping (side decoder), $P(\lambda_c)$ is the probability that λ_c is selected or $P(Q(X) = \lambda_c)$. In Shannon's information theory entropy is defined as a measure of disorder, or more precisely unpredictability of an information source. The entropy of a discrete random variable X with possible values of $\{x_1, x_2, \dots, x_n,\}$ is defined as (Shannon, 1948)

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_b p(x_i) \qquad (2\text{-}27)$$

where b is the base of the logarithm which determines the unit of entropy. If $b = 2$ then the unit is *bit*, and if $b = 10$ then the unit is called *dit*. The definition of the entropy for the discrete source can be extended to be used for continuous source, which is called differential entropy. The differential entropy for a continuous source with probability density function f_X is defined as (Shannon, 1948)

$$h(X) = \int_X f_X(x) \log_b f_X(x) \qquad (2\text{-}28)$$

If a single-description source X is blocked into L-dimensional vectors then the minimum entropy required to achieve an expected central distortion d_c is denoted as (Ostergaard et al., 2006)

$$R_c = \frac{H(Q(X))}{L} = -\frac{1}{L} \sum_{\lambda_c \in \Lambda_c} \int_{V_\Lambda(\lambda_c)} f_X(x)\, dx \log_2 \left(\int_{V_\Lambda(\lambda_c)} f_X(x)\, dx \right) \qquad (2\text{-}29)$$

Assuming that the Voronoi regions of all lattice points have identical areas of vol and the pdf is constant through each Voronoi region, R_c is approximated by