สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

A STUDY ON STEREOSCOPIC IMAGE ACQUISITION AND COMPRESSION

KUNTPONG WORARATPANYA

เลขหมู่.....
เลขทะเบียน.....46637
วัน,เดือน,ปี.....12 ก.ย. 2549

.b.....
.i.....

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
DOCTOR OF ENGINEERING IN ELECTRICAL ENGINEERING
SCHOOL OF GRADUATE STUDIES
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2005

ISBN 974-15-1836-6

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

COPYRIGHT 2005

SCHOOL OF GRADUATE STUDIES

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

หัวข้อวิทยานิพนธ์	การศึกษากระบวนการรับและการบีบอัดภาพสเตริโอสามมิติ
นักศึกษา	นายกันต์พงษ์ วรรัตน์ปัญญา
รหัสประจำตัว	41060018
ปริญญา	วิศวกรรมศาสตรดุษฎีบัณฑิต
สาขาวิชา	วิศวกรรมไฟฟ้า
พ.ศ.	2548
อาจารย์ผู้ควบคุมวิทยานิพนธ์	รองศาสตราจารย์ ดร.รัตติกร วรากูลศิริพันธุ์

บทคัดย่อ

ปัจจุบันเทคโนโลยีภาพและวิดีโอสเตริโอสามมิติถูกนำไปประยุกต์ใช้กับงานหลากหลายประเภท ซึ่งเทคโนโลยีประเภทนี้จะทำให้ผู้ชมมีปฏิสัมพันธ์ประหนึ่งว่าได้อยู่ในสถานการณ์นั้น อย่างไรก็ตาม ภาพและวิดีโอสเตริโอสามมิติดังกล่าวยังมีขนาดใหญ่ ซึ่งต้องการพื้นที่หน่วยของความจำจำนวนมากในการจัดเก็บข้อมูลและต้องการแบนด์วิดธ์ที่กว้างในการส่งข้อมูล ดังนั้นเพื่อเป็นการลดปริมาณและความซ้ำซ้อนของข้อมูลภาพและวิดีโอสเตริโอสามมิติลง การใช้ระบบการสร้างและเทคนิคการบีบอัดข้อมูลที่มีประสิทธิภาพจึงเป็นสิ่งจำเป็น วิทยานิพนธ์ฉบับนี้จึงได้นำเสนอระบบการสร้างภาพและวิดีโอสเตริโอสามมิติและเทคนิคการบีบอัดข้อมูลที่มีประสิทธิภาพ โดยระบบการสร้างภาพได้ใช้กลไกทางแสงซึ่งประกอบด้วยอุปกรณ์แบบแอคทีฟและแบบพาสซีฟ กลไกดังกล่าวจะทำงานตามมาตรฐานการบันทึกภาพแบบลำดับฟิลด์โดยใช้กล้อง CCD เพียงตัวเดียว สำหรับวัตถุประสงค์หลักของการวิจัยในส่วนนี้มีสองประการ คือ ประการที่หนึ่งเพื่อลดความซับซ้อนทางเรขาคณิตของการจัดวางตำแหน่งกล้องในการสร้างภาพสามมิติลง และประการที่สองเพื่อลดปริมาณสัญญาณของภาพและวิดีโอสเตริโอสามมิติลง ดังนั้นการศึกษาวิจัยในประเด็นของเรขาคณิตการจัดวางตำแหน่งกล้องจึงเป็นประเด็นที่สำคัญ เพื่อแทนที่เรขาคณิตของการจัดวางตำแหน่งกล้องแบบใช้กล้องสองกล้องด้วยเรขาคณิตการจัดวางตำแหน่งกล้องแบบใช้กล้องตัวเดียว วิธีการนี้ไม่เพียงแต่ระบบจะสามารถทำงานในเวลาจริงได้เท่านั้น แต่ยังสามารถลดความซับซ้อนของเรขาคณิตการวางตำแหน่งกล้องและลดปริมาณข้อมูลของภาพวิดีโอลงเมื่อเปรียบเทียบกับการใช้กล้องแบบสองกล้อง

แม้ว่าระบบการสร้างภาพและวิดีโอสเตริโอสามมิติด้วยกลไกดังกล่าว จะสามารถลดปริมาณข้อมูลของภาพวิดีโอลงได้ถึงครึ่งหนึ่ง แต่ก็ยังคงมีปริมาณที่มากอยู่ ด้วยเหตุนี้จึงได้นำเสนอเทคนิคการบีบอัดข้อมูลแบบแฟรคตอลร่วมโดยอาศัยการทำงานร่วมกับวิธีตรวจสอบความแปรปรวนซึ่งได้รวมจุดเด่นของทั้งสองวิธีเข้าด้วยกัน โดยมีวัตถุประสงค์ในการศึกษาวิจัยในส่วนนี้เพื่อให้องค์ประกอบหลักของการเข้ารหัสมีประสิทธิภาพทัดเทียมกัน องค์ประกอบเหล่านี้ได้แก่ คุณภาพของภาพ เวลาของการเข้ารหัส เวลาของการถอดรหัส และอัตราการบีบอัดข้อมูล วิธีการเข้า

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

รหัสแฟรคตอลแบบตรวจสอบความแปรปรวนเป็นวิธีการที่ง่ายและมีประสิทธิภาพในการเข้ารหัสที่รวดเร็วและให้คุณภาพของภาพที่ดี แต่ในส่วนของเวลาของการถอดรหัสและอัตราการบีบอัดข้อมูลก็ไม่ได้แตกต่างไปจากวิธีการดั้งเดิม ในอีกด้านหนึ่งถึงแม้วิธีการเข้ารหัสแฟรคตอลร่วมจะถูกพัฒนาในองค์ประกอบของเวลาของการถอดรหัสและอัตราการบีบอัดข้อมูลให้ดีขึ้นแล้วก็ตาม แต่ความเร็วของการเข้ารหัสก็ยังคงต้องพัฒนาให้เร็วขึ้น ดังนั้นเพื่อให้บรรลุวัตถุประสงค์ทั้งสี่องค์ประกอบข้างต้น การออกแบบที่พักสำหรับรวบรวมโดแมนแบบร่วมจึงถูกนำมาประยุกต์ใช้งานกับวิธีตรวจสอบความแปรปรวน ซึ่งผลการทดลองได้พิสูจน์ให้เห็นว่า วิธีการที่นำเสนอดังกล่าวสามารถทำให้การเข้ารหัสกระทำได้อย่างมีประสิทธิภาพ

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Thesis Title	A Study on Stereoscopic Image Acquisition and Compression
Student	Mr. Kuntpong WÒRARATPANYA
Student ID	41060018
Degree	Doctor of Engineering
Programme	Electrical Engineering
Year	2005
Thesis Advisor	Assoc. Prof. Dr. Ruttikorn VARAKULSIRIPUNTH

ABSTRACT

Stereoscopic image and video technology is being used increasingly in many applications. This technology provides the viewer with the feeling of "being present in the scene", however, stereoscopic image and video sequences require considerable storage capacity and transmission bandwidth. In order to minimize the signals of these sequences, therefore, effective acquisition systems and compression techniques are essential. In this dissertation, an effective acquisition system for generating the stereoscopic image and video sequences by using an active/passive optical adapter is proposed. This adapter, which is based on a field-sequential standard for stereoscopic image and video, collaborates with a single CCD camera. The primary objectives of this system are to decrease the geometrical complexities and streams of the stereoscopic image and video. The stereoscopic camera geometry is a principal issue that is investigated to substitute dual camera geometry with single camera geometry collaborating with the proposed adapter. In this way, not only does the system perform in real time, but it also reduces the geometrical complexities and signal streams of the stereoscopic image and video, when compared with the dual camera technique.

Although the proposed acquisition system can reduce the stereoscopic image and video sequences to be half of the whole, these sequences still have considerable information. For this reason, the mutual fractal coding based on variance is proposed. This approach is a combination of the advantages of mutual fractal coding and variance-based fractal coding. The main objective of the proposed method is to compromise the following factors: image quality, encoding time, decoding time and coding bit rate. The variance-based scheme is a simple and effective approach to improve fast encoding time and achieve good image quality, but the decoding time and coding bit rate are still no worse than the conventional fractal coding. On the other hand,

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

although the mutual fractal coding can provide both less decoding time and lower bit rate, the encoding time still has to speed up. Therefore, in order to accomplish the objective, mutual domain pools are designed and implemented for a variance-based scheme. In this way, the experimental results show that the proposed method can perform effectively.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Acknowledgements

I am very grateful to my research advisor, Assoc. Prof. Dr. Ruttikorn VARAKULSIRIPUNTH, for his guidance and support throughout my years at King Mongkut's Institute of Technology Ladkrabang. I would like to specially thank Prof. Dr. Yasushi KATO who kindly suggested valuable issues for my research work. I would also like to thank all members of my dissertation committee. Their recommendations absolutely improved the quality of this dissertation.

I would like to thank all members of the Communication Networks Laboratory, Research Center for Communications and Information Technology (ReCCIT): specially, Chompoonoot TENGCHAROEN, Jirapat KITTIMETHEE, Sumet PRABHAVAT, and Narissorn SANGKANONG for group discussions and technical supports.

I would like to thank Japan International Cooperation Agency (JICA) for supporting the research equipment.

I especially wish to thank Assoc. Prof. Dr. Monchai TEINTHONG, Dean of Faculty of Information Technology, and Assist. Prof. Dr. Supot NITSUWAT, Deputy Dean of Faculty of Information Technology, KMITNB, who have given me the opportunity to lecture a course of Information Theory at this faculty.

I would like to thank all colleagues at the Department of Computer Education, Faculty of Technical Education, KMITNB, who encourage me to complete my doctoral degree.

I wish to thank Natsaran PUTTHISOMBATKUL for her endless inspiration throughout my graduate studies.

I owe a great deal to my parents, brothers, and sisters. Although they are not involved with the technical aspects of this work, the warm emotion that they give me is largely responsible for completion of this work.

Kuntpong WORARATPANYA

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Table of Contents

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content and cite the document when use.

Table of Contents (Cont.)

This material is protected by copyright and is not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Table of Contents (Cont.)

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

List of Figures

Forbidden to modify the content, and cite the document when use.

List of Figures (Cont.)

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Chapter 1

Introduction

1.1 Significant Problems

The technology of three-dimensional (3-D) images and videos is currently becoming an important ingredient of multimedia applications [1], [2], [3], [4], [5], [6], [7], [8]. This technology provides the viewer with the feeling of "being present in the scene". The field of 3-D images and videos can be categorized into two main classes: stereoscopic image and video sequences and multiview image and video sequences [9]. These image and video sequences can also be generated by computer-generated [9], [10], [11] and real-imaging capture approaches. Here, only the field of the stereoscopic image and video sequences formed by real-imaging capture approach are studied, since these sequences are made up of the imitation of the human-binocular vision system, which is a challenge field of research [4], [5], [6], [7].

Stereoscopic image and video sequences require considerable storage capacity and transmission bandwidth, therefore, in order to minimize the signals of the stereoscopic image and video sequences, effective acquisition systems and compression techniques are essential. In this dissertation, two collaborating systems, acquisition and compression, are investigated for reducing the redundancy of stereoscopic image and video sequences. For the acquisition system, the stereoscopic image and video format and camera configuration for encoding the stereoscopic image and video sequences are very important. In case of format standards, there are at least five principal methodologies by which stereoscopic image and video sequences can be encoded into standard-video signals, such as field-sequential, side-fields, sub-fields, separated channels, and anaglyph standards [12]. The *field-sequential standard* is one of the most practical approaches for producing and displaying stereoscopic image and video sequences on existing equipment, such as computer monitors and televisions [2]. It can be also applied directly to other types of stereoscopic displays, such as anaglyphic displays, polarized projected displays, half-silvered mirror displays and some lenticular displays [13]. In the case of the camera, there are two types of configurations, i.e., toed-in camera configuration and parallel camera configuration. However, the former causes keystone distortion, while the latter does not exhibit this

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

distortion [13]. For this reason, this dissertation works with the *field-sequential standard* and *parallel camera configuration*.

For a compression system, the field-sequential stereoscopic image and video sequences are generally composed of left and right images acquired from two slightly different viewpoints, thus making them similar and containing a lot of redundant information. In order to remove this redundancy, an effective coding technique is essential. With a characteristic of the field-sequential stereoscopic image and video sequences containing a lot of self-similarity information, the suitable coding approach for such information is *fractal coding*, based on the self-similarity property. This technique has various advantages in terms of resolution independence, fast decoding and high compression ratio, however, long encoding time is still a main drawback [14]. In the case of a frame-based image coding of the field-sequential stereoscopic video sequences, the conventional 2-D fractal coding technique makes the depth or three-dimensional perception defective, since the high vertical-frequencies of a stereoscopic image pair are suppressed [15]. Therefore, the proposed coding system has to compromise the following factors: image quality (including depth perception), encoding time, decoding time and coding bit rate.

1.2 Prior Researches

In this section, prior works related to the problems of stereoscopic video acquisition and stereoscopic image coding are reviewed. The advantages and disadvantages of various techniques and methods are also presented as follows:

1.2.1 Stereoscopic Image and Video Acquisition

Prior works of stereoscopic video acquisition, in order to present a viewer with stereoscopic images or videos, made use of various methods of image separation and recombination, including dual camera configuration and single camera configuration. The dual camera configuration usually makes the storage and bandwidth double, therefore, in this dissertation, only a single camera configuration is considered and reviewed.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

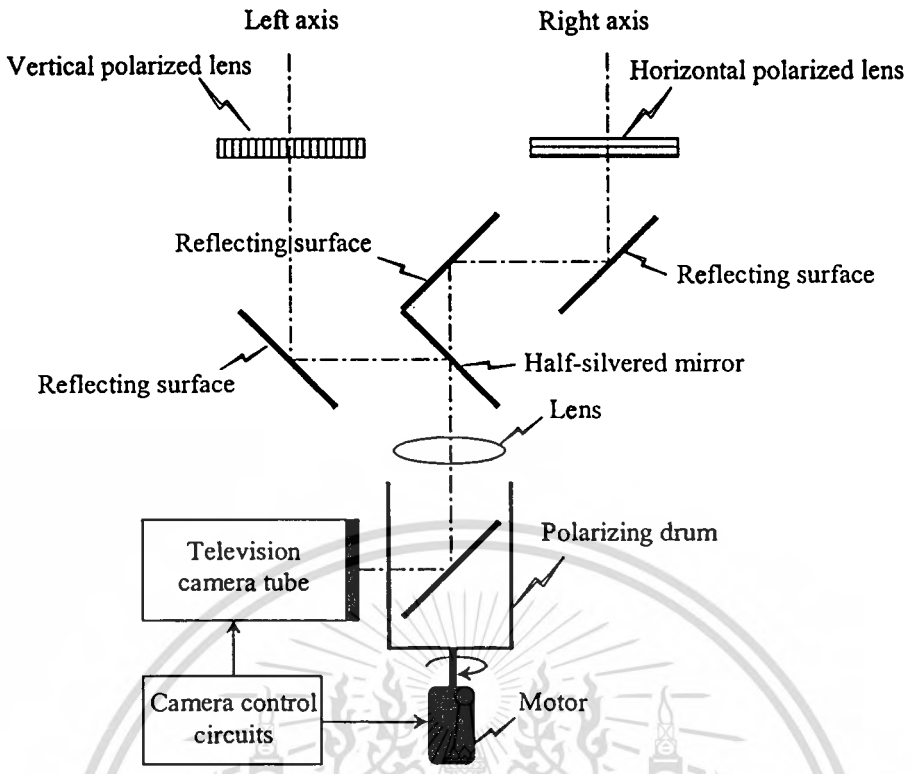


Figure 1.1 Stereoscopic image acquisition with a single camera tube and optical system.

The 1950 U.S. Pat. No. to R. D. Kell, 2,508,920 [16] discloses that a television system can encode stereoscopic images to a standard television transmission in a field-sequential mode. The camera configuration is schematically depicted in Figure 1.1. The optical system comprises a lens for focusing the images to a light responsive electrode surface. In the front of the lens, there is a half-silvered mirror and in the front of this mirror, there is a reflecting surface. The reflecting surfaces are also provided on the left and right hand side of the half-silvered mirror. In the front of these reflecting surfaces, there is a vertical polarized lens placed on the left axis, while a horizontal polarized lens placed on the right axis, respectively. In this alignment, the camera configuration is very interesting, since the optical system makes use of fewer optical elements. However, the heart of this system is a vertical and horizontal polarized drum controlled by a motor for encoding the left and right images into the field-sequential mode. The motor is a fluctuation mechanic and its size is ordinarily larger. Furthermore, the use of a half-silvered mirror in the optical system leads to more than 50% loss of light reaching the light responsive electrode surface [17]. These are the disadvantage of this optical system.

This thesis is available for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

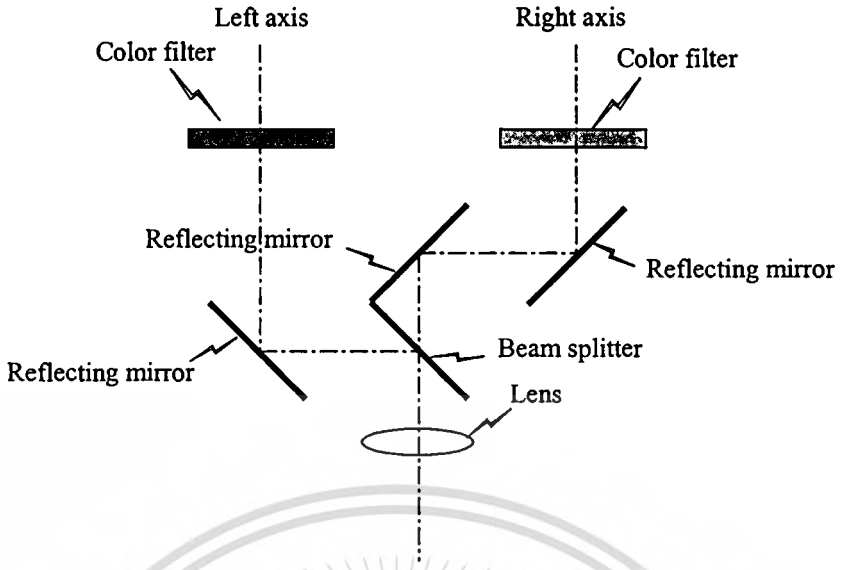


Figure 1.2 Stereo-optical device attached to a lens of a three-tube color camera.

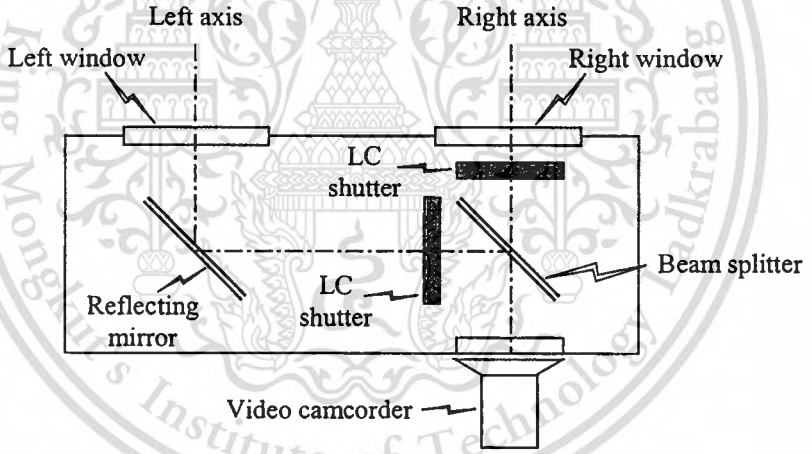


Figure 1.3 Stereoscopic converter assembly.

The 1975 U.S. Pat. No. to Roese, 3,903,358 [18] discloses that a PLZT stereoscopic television system functions to present alternate left and right images to a single video whose output is fed to a television monitor. The camera configuration is similar to the one proposed by D. R. Kell, but Roese's system concentrates on the improvement of material by replacing the polarized lenses with ceramic light valves (lanthanum modified lead zirconate-tinanate: PLZT).

The 1988 U.S. Pat. No. to J. F. Butterfield, 4,734,756 [17] discloses a stereoscopic acquisition system. In this system, schematically illustrated in Figure 1.2, a

Forbidden to modify the content, and cite the document when use.

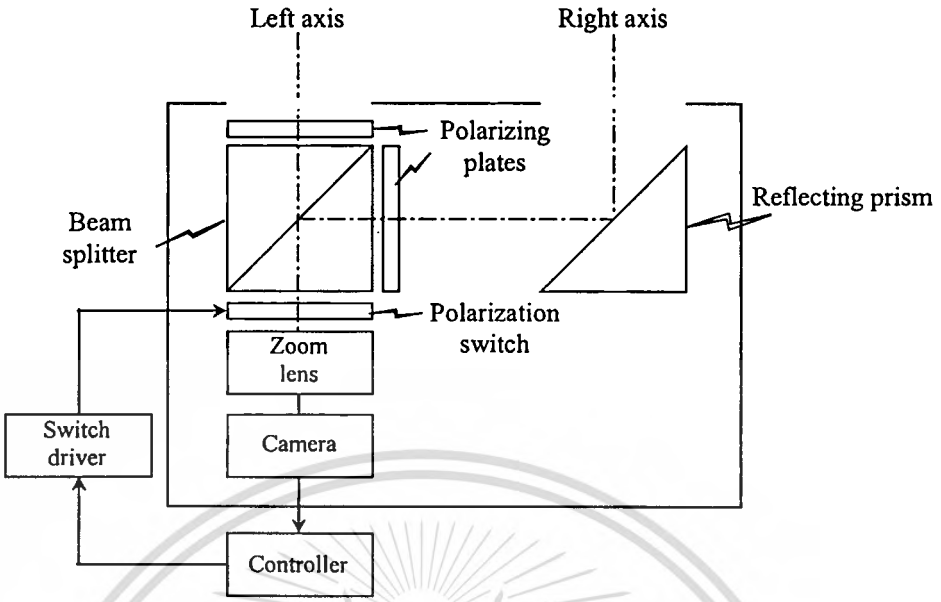


Figure 1.4 Alignment of optical devices for stereoscopic image acquisition.

stereo-optical device, attached to a lens of a three-tube color camera, is comprised of a pair of inner mirrors and outer mirrors. A number of mirror arrangements are similar to the one proposed by R. D. Kell. Nonetheless, one of inner mirrors is half-silvered. This leads to light reduction. Furthermore, a beam-splitter, used for reflecting light from the single lens to three tube color cameras (omitted in Figure 1.2), generates a ghosting image due to the reflection from the second beam-splitter surface [19]. This ghosting image can lead to headaches [5].

The 1990 U.S. Pat. No. to P. A. Femano, 4,943,852 [20] proposed a stereoscopic converter assembly. The optical-device alignment of the converter assembly is illustrated in Figure 1.3. The significant advantages of this configuration are simple geometry and fewer optical devices. On the other hand, the beam-splitter may generate a ghosting image [19]

The 1991 U.S. Pat. No. to Sudo, 5,003,385 [21] discloses the alignment of optical devices for acquiring stereoscopic images as shown in Figure 1.4. The optical device configuration is close to the one proposed by P. A. Femano, but the principal difference is that the prisms are exploited in this configuration. On the viewer side, the monitor screen has to be covered by a polarization switch.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

As preceding overviews of evolution of the stereoscopic image and video acquisition system, almost principal optical-device alignments for an acquisition system have no great difference, only the types of material are changed so as to improve the quality of images and videos.

1.2.2 Stereoscopic Image Coding

Stereoscopic image sequences are acquired from two slightly different viewpoints, thus making them similar and containing a lot of redundancy. Independent coding of these stereoscopic image pairs may lead to be inefficient. Therefore, H. Yamaguchi [22] investigated the statistical characteristics of stereoscopic image disparity for removing the stereoscopic image redundancy. In 1992, the stereoscopic image compression exploiting the correlation between stereoscopic images was proposed by M.G. Perkins [3]. Since then, many researches have been proposed on various techniques for coding stereoscopic image pairs as reported in literatures [1], [23], [24], [25], [26]. The stereoscopic image coding based on the disparity compensation (DC) is the most favored approach, especially, the disparity field of a stereoscopic image pair obtained by *intensity-based* method. This approach can be briefly reviewed as follows. The left image is encoded independently (separated from the right image) and transmitted to the receiver. Then the right image is divided into sub-image non-overlapping blocks. The disparity of each sub-image block in the right image is estimated and sent to the receiver. Finally, the residual image of each sub-image block in the right image is predicted and also transmitted to the receiver. In this method, the heart of coding is the disparity estimation process. This process is closely related to motion estimation, widely used in the 2-D video codecs. However, there are subtle differences between DC and MC as listed below [1], [3], [26]:

- The disparity d consists of two components: a horizontal and a vertical component, d_x and d_y, respectively. When the stereoscopic camera configuration is aligned in parallel axes geometry, d_y is always zero and d_x is positive, whereas in the 2-D video sequences, the object in an image can move on vertical and horizontal displacements in both directions.
- In a stereoscopic image pair, every object in the left image is displaced in the right image, while only a few objects change their locations from frame to frame in a typical video sequence.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

- In a video sequence, the intensities of corresponding in sequential frame can be assumed to be equal. However, this assumption is often false in a stereoscopic image pair, because of individual differences between the left and right cameras.

Many papers, [1], [3], [26] have concluded that one of the drawbacks of stereoscopic image coding based on disparity compensated prediction is occluded regions, which cannot be represented by the reference image (the left image) and disparity vectors alone. At this point, it was found that the stereoscopic image coding techniques have been constantly developing.

1.3 Objectives of Research

For the acquisition system, the primary objectives are to decrease the geometrical complexities and the streams of stereoscopic video and also show that the dual camera stereoscopic system can be replaced with the proposed acquisition adapter. Furthermore, the signal streams of stereoscopic video acquired by the proposed approach are compatible with the field-sequential stereoscopic standard.

For the compression system, the significant purpose is to investigate a novel concept of fractal coding for stereoscopic image pairs. In addition, the theoretical model of this concept proves that it can be applied to stereoscopic image pairs in both cases: frame-based and field-based images. The proposed coding system also has to compromise the following factors: image quality, encoding time, decoding time and coding bit rate.

1.4 Limitations

In this dissertation, there are two limitations for the research work:

(i) Intra-frame (including intra-field) is a key frame in a video sequence and the overall coding efficiency depends on this frame, thus, the intra-frame coding is a main issue of the research work.

(ii) A comparison of proposed schemes or methods and the other standards is not a significant issue for this research.

1.5 Approach

In this section, the approaches to achieve the objectives set in Section 1.3 are outlined. In the following subsections, the concepts of an acquisition adapter and coding approach for stereoscopic images are presented.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

1.5.1 Acquisition Adapter

In this dissertation, an acquisition adapter is created by the following conditions. First, the adapter has to be compatible with a field-sequential standard, which is the most practical approach for producing and displaying stereoscopic image sequences on existing equipment. This standard can be also applied directly to other types of stereoscopic displays. Second, this adapter has to be compatible with a parallel camera configuration, which does not cause keystone distortion. Finally, the signal streams of the stereoscopic video acquired by this adapter have to be minimized, therefore an active/passive optical adapter has been created as described in details in Chapter 2.

1.5.2 Proposed Coding Approach

In this subsection, we begin with analyzing a conventional scheme of transform coding for stereoscopic image pairs. As presented in Subsection 1.2.2, the heart of the transform coding scheme shown in Figure 1.5 is disparity estimation and residual image coding. Most research focuses on these processes as reported in [3], [23], [24], [25], [26], [27]. This is because motion estimation and discrete cosine transform (DCT), which are widely used for typical videos and images, cannot be applied directly to disparity estimation and residual image, respectively. The disparity parameters and residual-image-transformed coefficients are side information that has to be sent to receiver for recovering the right image. The main drawback of the disparity estimation process is occluded regions; the portion of image appears in the left image, but disappears in the right image and vice versa. In addition, the shortcoming of residual image coding is that the existing transform coding does not always achieve a good coding efficiency [26]. Thus, the residual image needs a special design of transform coding scheme as proposed in [23], [24], [27].

In this dissertation, an alternative way is investigated so as to overcome these problems by using fractal coding, especially, affine transformation (commonly used in fractal image compression) defined by

$$F_i\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a_i & b_i & 0 \\ c_i & d_i & 0 \\ 0 & 0 & s_i \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} e_i \\ f_i \\ o_i \end{bmatrix} \tag{1.1}$$

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

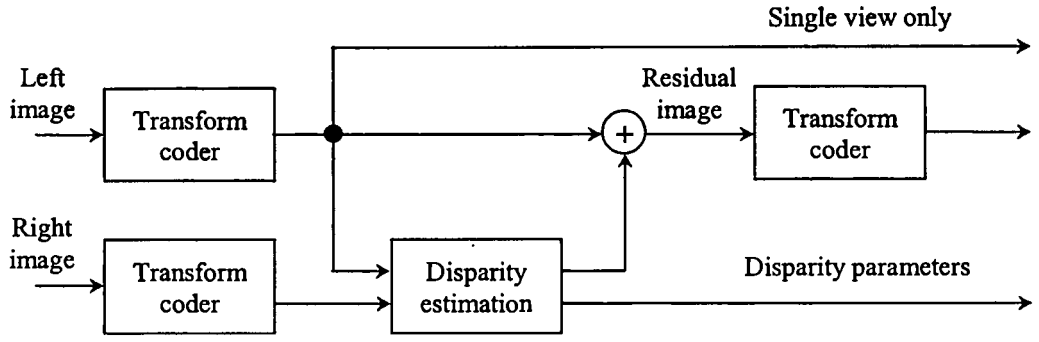


Figure 1.5 Transform coding scheme for stereoscopic image pairs.

builds in geometry and intensity parts that can represent the disparity parameters and residual-image-transformed coefficients, respectively. In general, Eq.(1.1) contains eight parameters *a*, *b*, *c*, *d*, *e*, *f*, *s* and *o*, where the first six-parameters (*a*, *b*, *c*, *d*, *e* and *f*) denote coefficients of geometry transform, while the others (*o* and *s*) represent coefficients of intensity transform. In the disparity estimation process, a sub-block of right image (target) is searched for the matching block in the left image (reference) by means of mean absolute difference (MAD) or mean-square error (MSE) [28]. In this phase, the positions of all corresponding blocks of the right image are retained as a disparity vector field. The operations of this step are close to that of the domain-range mapping procedure for encoding the fractal codes. In addition, in the case of disparity-compensated transform-domain predictive coding (DCTDP) proposed by M.G. Perkins, the estimation of the transform of the right image sub-blocks from the transform of the matching left image sub-block makes use of a linear predictor as defined by

$$\hat{R} = X \otimes L + Y, \tag{1.2}$$

where operator $\otimes$ is the array multiplication; i.e., the $(i,j)^{th}$ element of the first matrix is multiplied by the $(i,j)^{th}$ element of the second matrix. The matrices X and Y are based on the statistics obtained from a training set, so that each coefficient is predicted by a minimum variance predictor. In this case, *o* and *s* parameters in affine transform provide similarly to the operation of the procedure that the DCTDP does.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

At this point, it is found that the appropriate design of fractal coding can perform the same way as the transform-coding scheme for stereoscopic image pairs. The disparity estimation process is omitted and the occluded region is not an obstacle. Consequently, the mutual fractal coding for stereoscopic image pairs is proposed. The concept of mutual fractal coding is introduced in Chapter 3. Its extension is also described in Chapter 4. Finally, the combination of mutual fractal coding and variance-based approach is proposed in Chapter 5.

1.6 Dissertation Overview

This dissertation is organized as follows. In Chapter 2, the proposal is for a new approach for acquiring stereoscopic video by using an active/passive optical adapter. This adapter, which is based on a field-sequential standard for encoding stereoscopic-video format, collaborates with only a single CCD camera. The core of this research is to substitute the dual camera geometry with single camera geometry collaborating with the proposed adapter. In this way, not only does the system perform in real time, but it also reduces the geometrical complexities and signal streams of stereoscopic video, when compared with a dual camera technique.

In Chapter 3, the proposal is for the novel concept of fractal coding, i.e., mutual fractal coding. In this approach, there is a significant characteristic that differs from the conventional fractal coding. That is the effective domain pool design; the mutual domain pools based on the *equivalent* property of two complete metric spaces are designed for speeding up the encoding and decoding time. For applications to field-sequential stereoscopic image sequences, the remaining problem is that the high vertical-frequencies of stereoscopic image pairs in frame-based images are suppressed. This makes the depth perception defective. In this case, there are two simple techniques for overcoming this problem. The former technique is mutual scanline fractal coding (MSFC) for frame-based images. The latter is mutual 2-D fractal coding (M2DFC) for field-based images. It is an extended concept from the MSFC approach described in Chapter 4. The intended approaches can reduce both encoding and decoding times when compared with the conventional fractal coding based on quadtree partition. The reconstructed image quality is also improved in terms of preserving the high vertical-frequencies by using the MSFC for frame-based images.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

The mutual scanline fractal coding for frame-based images is successful not only for reducing the encoding and decoding time, but also for preserving the high-vertical frequencies, however, its coding efficiency is not satisfactory for very-low-bit-rate situation. Therefore, in Chapter 4 a two-dimensional approach is proposed, so that the coding efficiency in terms of bit rate is improved. In addition, the two-dimensional approach also provides the generic structure of coding images.

In Chapter 5, the mutual fractal coding based on variance is proposed. This approach is a combination of the advantages of the mutual fractal coding and variance-based fractal coding. The main objective of the intended method is to compromise the following factors: encoding time, decoding time, bit rate and image quality. The state-of-the-art of the variance-based scheme is a simple and effective approach to improve fast encoding time and to achieve good image quality, but the decoding time and bit rate are still no worse than the conventional fractal coding. On the other hand, although the mutual fractal coding introduced in Chapters 3 and 4 can provide both less decoding time and lower bit rate, the encoding time still has to be sped up. Therefore, in order to accomplish the objective, a mutual domain pool design is implemented for the variance-based scheme.

Finally, the conclusions and suggestions are given in Chapter 6. In this chapter, the advantages and disadvantages of acquisition and coding systems are discussed. Feasible extensions of this research work are also introduced.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Chapter 2

Stereoscopic Video Acquisition System

2.1 Introduction

The camera geometry for encoding stereoscopic videos is one of the main issues in the technologies of three-dimensional (3-D) acquisition systems, which have been rapidly developed and applied to many applications, such as robot vision, virtual machine, medical surgery, and teleconference. In general, stereoscopic images and videos are generated from two cameras placed in slightly different positions in a horizontal orientation. In this case, there are two types of camera configurations, i.e., toed-in camera configuration and parallel camera configuration. The former causes keystone distortion, while the latter does not exhibit this distortion as reported in [13]. For this reason, the parallel camera configuration (or parallel axes geometry) is investigated so as to design a stereoscopic video acquisition system. This configuration requires the following properties: The left camera's focal ray is parallel to the right camera's focal ray; the left camera's image plane and the right camera's image plane are coplanar; and, the bottom edge of the left camera's area and the bottom edge of the right camera's area are collinear [3]. These properties make the system difficult to set up. In addition, the video streams acquired by two cameras are double. The widening information is another obstacle, since it requires considerable storage space and transmission bandwidth. The previous restrictions are important problems that have been attempted to overcome. In this chapter, the principal objectives are not only to overcome the widening information, but also to provide an inexpensive and practical means for the viewers with true three-dimensional images and videos and with little or no modification to exiting equipment. Consequently, we concentrate on the issue of replacing dual camera geometry with single camera geometry, collaborating with an active/passive optical adapter (APOA).

The rest of this chapter is organized as follows: In Section 2.2, we introduce dual camera geometry and also propose an APOA adapter cooperating with a single camera. Then we analyze the geometry of the APOA adapter and test it, so as to verify the compatibility with the dual camera geometry. In Section 2.3, we describe

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

the method of encoding stereoscopic video sequences. Section 2.4 illustrates the experimental results and Section 2.5 is the conclusion.

2.2 Design of Stereoscopic Video Acquisition

In Chapter 1, some prior works related with acquisition systems are presented. The disadvantages of these systems are considered. Therefore, the main study in this section is the issue of replacing dual camera geometry with single camera geometry collaborating with the APOA adapter. This section comprises three subsections: geometry of dual cameras, geometry of active/passive optical adapter and analysis of stereoscopic camera geometry.

2.2.1 Geometry of Dual Cameras

In this subsection, we begin with describing the geometry of stereoscopic images and videos. Essentially, the construction is founded from two cameras as depicted in Figure 2.1. Each camera is composed of a simple lens and an image sensor. The left and right cameras are placed in two slightly different positions in a horizontal orientation. The distance B between left and right optical axes is called the *baseline*. This baseline is usually approximate to human inter-pupillary distance. The focus f is the distance between camera's lens and camera's image sensor. Based on parallel axes geometry, the geometrical configurations require the following properties: (i) The left and right optical axes must be parallel, (ii) the bottom edges of left and right images are collinear, (iii) the left and right image sensors are coplanar, and (iv) the left and right cameras must be identical. In practice, these properties are difficulties for setting up the system [3].

2.2.2 Geometry of Active/Passive Optical Adapter

According to the restrictions of dual camera geometry depending on parallel axes geometry, a single camera geometry is considered to substitute for the double camera geometry. However, only a single camera cannot generate and encode stereoscopic images and videos. The APOA adapter for collaborating with a single camera is necessary. The simple structure of the APOA's geometry depicted in Figure 2.2 consists of four important passive optical devices: a left prism, right prism, reflecting prism and right-angled prism [30], two active optical devices: left and right LC shutters and a CCD camera containing a simple lens and an image sensor. The

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

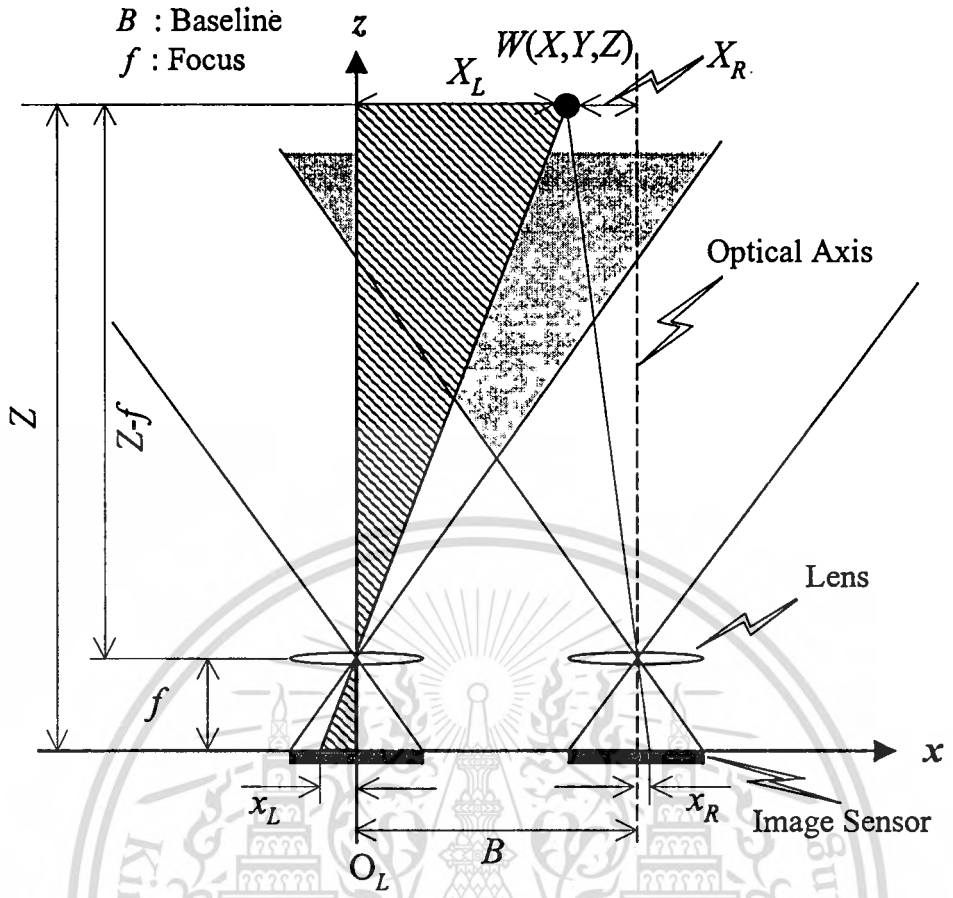


Figure 2.1 Top view of dual camera geometry.

parameters B and f are still similar to the dual camera geometry, but the left and right image sensors of the dual camera geometry are replaced with the left prism and the right prism, respectively. The main function of these prisms is to deviate or displace the optical axes illustrated it with dash lines in Figure 2.2. Thus, the left and right images entering to the left and right prisms are refracted by 90 degrees from the left to the right-angle prism and from the right to the reflecting prism, respectively. Finally, both images are projected to the CCD's image sensor by the right-angle prism and reflecting prisms. At this time, the CCD camera can grab and encode the image sequences coming from two slightly different perspectives into the stereoscopic-video signals described in details in Subsection 2.3.

In summary, the intended adapter not only eliminates the difficulty of setting up the system and reduces the signal streams, but also provides identical system parameters, such as blurring, lens distortion, focal length, spectral response, gain, offset, pixel size, etc [31].

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

2.2.3 Analysis of Stereoscopic Camera Geometry

From subsection 2.2.1 and 2.2.2, the components of dual camera geometry and single camera geometry collaborating with the APOA are formed and depicted in Figures 1 and 2, respectively. In this subsection, the camera geometry is analyzed and shows that the dual camera geometry can be completely replaced with the single camera geometry collaborating with the APOA. A ranging equation [32] is used to prove the compatibility of the dual camera geometry and the APOA with the single CCD. The equations of parallel axes geometry derived from Figure 2.1 are

$$\frac{-x_L}{f} = \frac{X_L}{Z-f};\ X_L = \frac{x_L}{f}(f-Z) \tag{2.1}$$

and

$$\frac{x_R}{f} = \frac{-X_R}{Z-f};\ X_R = \frac{x_R}{f}(f-Z)\,, \tag{2.2}$$

which are determined by using similar triangles. In this case, y_L and y_R are equal. The minus sign of x_L in Eq. (2.1) indicates that the image point, W, is actually inverted when the origin of coordinate system is set to the center of the image sensor of the left camera, O_L. The minus sign appears in the front of X_R in Eq. (2.2). To simplify the Eqs. (2.1) and (2.2), the relationship of X_L and X_R is assigned to $X_R + X_L = B$. Then, we can obtain the classic ranging equation as follow

$$Z = f\left(1 - \frac{B}{x_R - x_L}\right) = f\left(1 + \frac{B}{x_L - x_R}\right). \tag{2.3}$$

In the case of single camera geometry collaborating with the APOA, if the passive optical devices perfectly refract the point image W_S, which is equal to W, both left and right optical axes of prisms can be translated to the optical axis of CCD's lens. Hence, x'_{LS} and x'_{RS} can be written as:

$$x'_{LS} = x_{LS} + \frac{B}{2};\ X'_{LS} = X_{LS} + \frac{B}{2} \tag{2.4}$$

and

$$x'_{RS} = x_{RS} - \frac{B}{2};\ X'_{RS} = X_{RS} - \frac{B}{2}\,. \tag{2.5}$$

After translation, the ratio of similar triangles can be obtained by

$$\frac{-x'_{LS}}{a_1} = \frac{X_{LS} + B/2}{Z - a_1};\ \frac{x'_{LS}}{a_1} = \frac{X_{LS} + B/2}{a_1 - Z}\,, \tag{2.6}$$

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

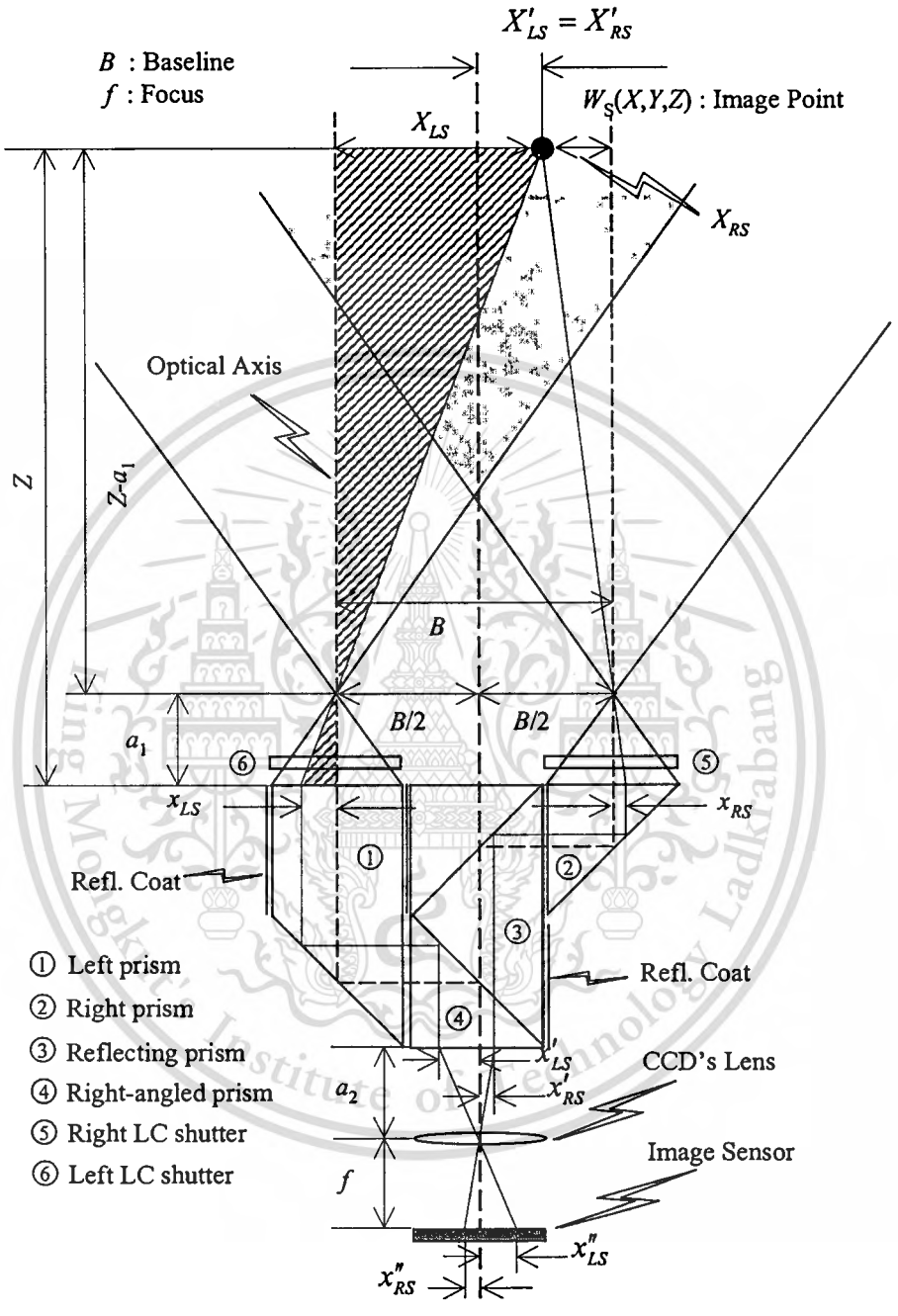


Figure 2.2 Top view of single camera geometry collaborating with APOA.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

and $$\frac{x'_{RS}}{a_1} = \frac{-(X_{RS} - B/2)}{Z - a_1}; \frac{x'_{RS}}{a_1} = \frac{X_{RS} - B/2}{a_1 - Z} \quad . \tag{2.7}$$

From Eqs. (2.6) and (2.7), to simplify them, the relationship of X'_{LS} and X'_{RS} is set to $X'_{LS} = X'_{RS}$. Then, we have the final result as

$$Z = a_1\left(1 - \frac{B}{x'_{RS} - x'_{LS}}\right) = a_1\left(1 + \frac{B}{x'_{LS} - x'_{RS}}\right) \tag{2.8}$$

where $x'_{LS} = -\frac{a_2}{f}x''_{LS}$ and $x'_{RS} = -\frac{a_2}{f}x''_{RS}$.

Eq. (2.8) shows that the ranging equation obtained from single camera geometry collaborating with the APOA is equivalent to Eq. (2.3), where $f, x_L,$ and x_R are similar to $a_1, x'_{LS},$ and x'_{RS}, respectively. Hence, the single camera geometry collaborating with the APOA is also comparable with the double camera geometry.

2.3 Implementation of Adapter for Encoding Stereoscopic Videos

Before discussing details of encoding stereoscopic videos, the background related to stereoscopic video standards is reviewed. There are at least five principal methodologies by which stereoscopic image sequences can be encoded into standard-video signals, such as field-sequential, side-fields, sub-fields, separated channels and anaglyph standards [12], but the most appropriate technique for this research is the field-sequential standard. This section only discusses the field-sequential stereoscopic video standard and its implementation.

2.3.1 Field-Sequential Stereoscopic Video Standard

Standards for implementing stereoscopic videos have many methods. The primary reason why the field-sequential method is widely exploited in stereoscopic videos is to handle worldwide televisions and video recorders without modifying any equipment. It can be also applied directly to other types of stereoscopic displays such as anaglyphic displays, polarized projected displays, half-silvered mirror displays and some lenticular displays [13]. The construction of a field-sequential stereoscopic standard is illustrated in Figure 2.3. This structure is compatible with the standard

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.
46637

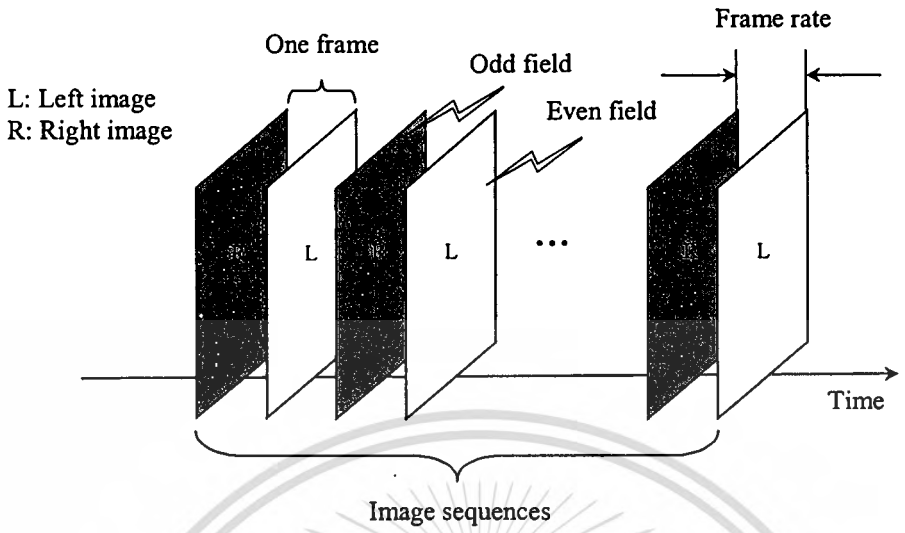


Figure 2.3 Structure of field-sequential stereoscopic standard.

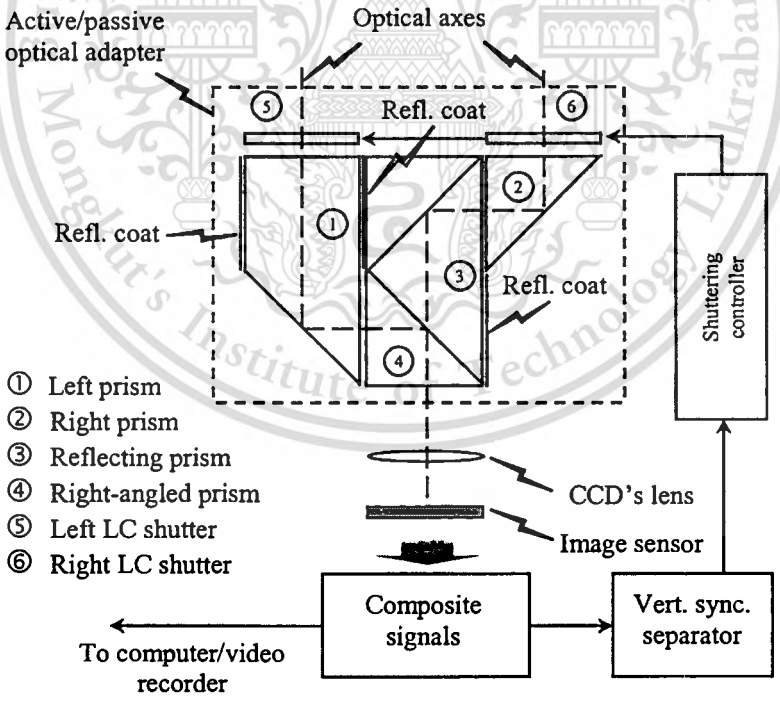


Figure 2.4 Model of stereoscopic video acquisition system using active/passive optical adapter.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

scanning of television systems, for which every frame is made up of two fields, even and odd. The number of times for repainting the screen in one second is called the *field rate*. This field rate is customarily related to a vertical synchronous frequency, 60 Hz for a NTSC system. With this field rate, it is expeditious that the human eye cannot identify the alternation between even and odd fields. Thus, the field-sequential method is a pertinent technique for occupying the left and right images in one frame, consisting of an even field for the left image and an odd field for the right image and for displaying on standard televisions or computer monitors.

2.3.2 Implementation of Stereoscopic Video Acquisition System

Implementation is the final procedure for this chapter. In Figure 2.4, the diagram of encoding stereoscopic videos by using a single camera collaborating with the APOA is manifested. The system is composed of a set of passive optical devices, a pair of LC shutters, a vertical synchronous separator, a shuttering controller and a CCD camera. The important part for encoding the left and right images into the even and odd fields of standard-video signals is a pair of LC shutters. The LC shutter is capable of transmitting and blocking light by controlling the voltage applied to the liquid crystal and has a sufficiently fast response to the field scanning frequency of the CCD camera. It also has other advantages of longer life and easier handling [33].

For encoding the left image, the manipulation of the proposed scheme commences from a light ray passing through the left LC shutter when it is turned on; it is transparent, while the right LC shutter is turned off; it is opaque, and then striking to the image sensor. At this time, the image sensor grabs the left image and encodes it into the even field of video signals. For encoding the right image, the operation is the same as mentioned previously. The dispatch rate of a pair of LC shutters is equal to field rate of vertical synchronous frequency. In this way the standard-video signals become stereoscopic-video signals.

The concept of observing stereoscopic videos encoded into a field-sequential standard is to separate the left-eye viewing for left image and the right-eye viewing for right image independently for producing depth (three-dimensional) perception [29], [33]. In the case of exhibiting stereoscopic videos on standard televisions or computer monitors, the viewers must wear a pair of LC shuttering glasses for watching depth information. These glasses aid them to be able to watch the left image by the left eye and the right image by the right eye separately. The function of LC

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

shuttering glasses is as follows: When the even field keeping the left image is presented on the screen, the left side of the LC shuttering glasses is transparent and the opposite side is opaque. Similarly, when the odd field holding the right image appears, the right side of the LC shuttering glasses is transparent and the opposite side is opaque. Then the left and right images are processed with the human brain and the depth perceptions are created for the viewers' eyes.

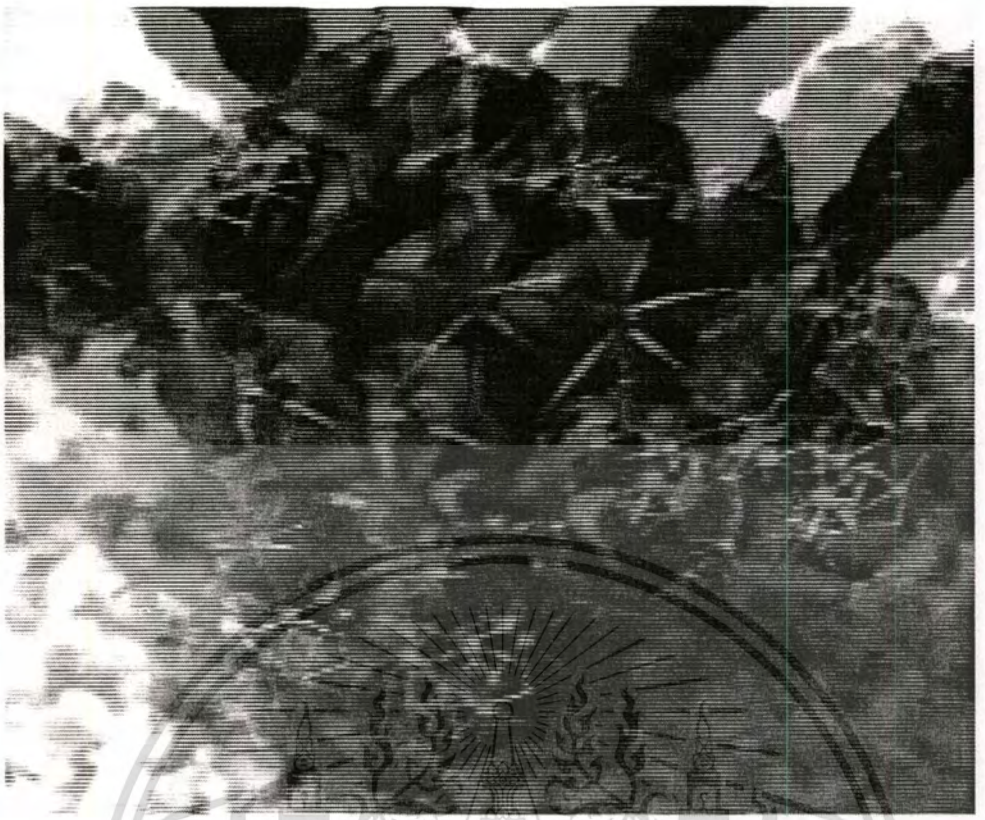
2.4 Experimental Results

Figure 2.5 illustrates the experimental results, a set of stereoscopic image pairs generated from the prototype system and a single camera collaborating with the APOA. For this experiment, the parameter B is set to 50 mm. Figure 2.5 also shows that the frame holds both left image in an even field and right image in an odd field. When viewers see this picture with the naked eye on a screen, television or computer monitor operating on interlaced-scanning system, the left and right images appeared in the condition of transparent overlap, as if they were displayed at the same time. In fact, they are alternately exhibited field-by-field. On the other hand, if the viewers look through a pair of LC shuttering glasses synchronized to field rate of vertical synchronous frequency for each eye, they can observe the depth perspective video.

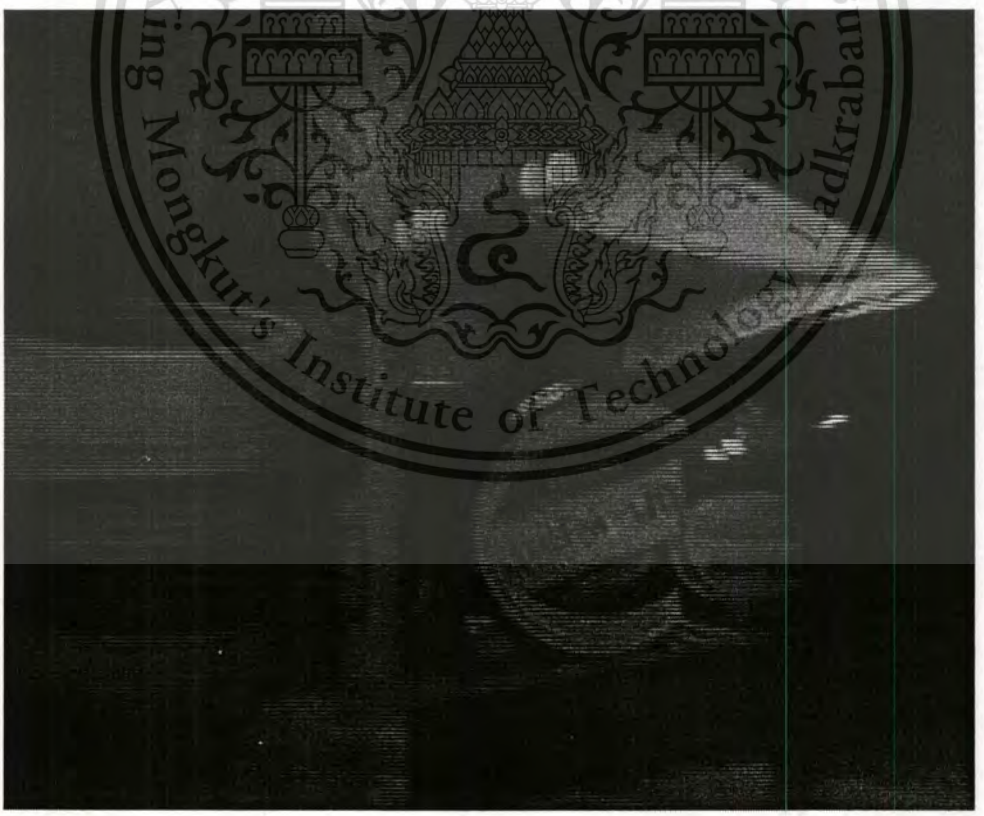
2.5 Conclusion

The acquisition adapter for encoding stereoscopic videos utilizing single camera was introduced and also demonstrated the outcome from a prototype system. By this approach, not only can the system perform in real time, but it can reduce the number of cameras, the number of video streams and the camera geometrical complexities as well. In the case of transmitting stereoscopic video into communication channels, the stereoscopic video stream generated by the single camera collaborating with the APOA is less than by dual cameras two times. Although the objective was achieved, a few limitations still remain. For instance, the system cannot magnify image sequences when taking objects that are far away and the resolution of images is cut down by half, because of the characteristics of the field-sequential standard. In case of stereoscopic video quality, it is dependent on the performance of the CCD camera. In next chapter, the redundancy of field-sequential stereoscopic video sequences is reduced by means of an effective coding.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



(a)



(b)

Figure 2.5 Stereoscopic image pairs generated from the archetype acquisition system.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Chapter 3

Concept of Mutual Fractal Coding: A One-dimensional Approach

3.1 Introduction

Stereoscopic image sequence is becoming an important ingredient of multi-media applications. The field-sequential stereoscopic image sequence especially is one of the most practical approaches for producing and displaying on existing equipment, nevertheless, it requires an effective coding to minimize its information for storage space and transmission bandwidth. In this chapter, the effective coding technique and the characteristics of the field-sequential stereoscopic image are collaboratively investigated.

The field-sequential stereoscopic image sequences are composed of left and right images acquired from two slightly different viewpoints, thus making them similar and containing a lot of redundant information. In order to remove this redundancy, an effective coding technique is essential. With a characteristic of the field-sequential stereoscopic image sequences containing a lot of self-similarity information, the suitable coding approach for such information is fractal coding based on the self-similarity property. This technique has various advantages in terms of resolution independence, fast decoding and high compression ratio (or very low bit rate), however, long encoding time is still a main drawback. In case of frame-based image coding of the field-sequential stereoscopic image sequences, the conventional 2-D fractal coding technique makes the depth or three-dimensional perception defective, since high vertical-frequencies of a stereoscopic image pair are suppressed.

In order to overcome these problems, a novel coding approach, mutual fractal coding, is proposed for coding the field-sequential stereoscopic image sequences. In this approach, there is a significant characteristic that differs from the design of the conventional fractal coding; that is the effective domain pool design. It is one of the important phases for speeding up the encoding time of fractal coding. A. E. Jacquin [34] concluded his unsolved problems concerning the concept of fractal encoding. The domain pool is one of his lists. Since then, fewer papers [35] have been proposed

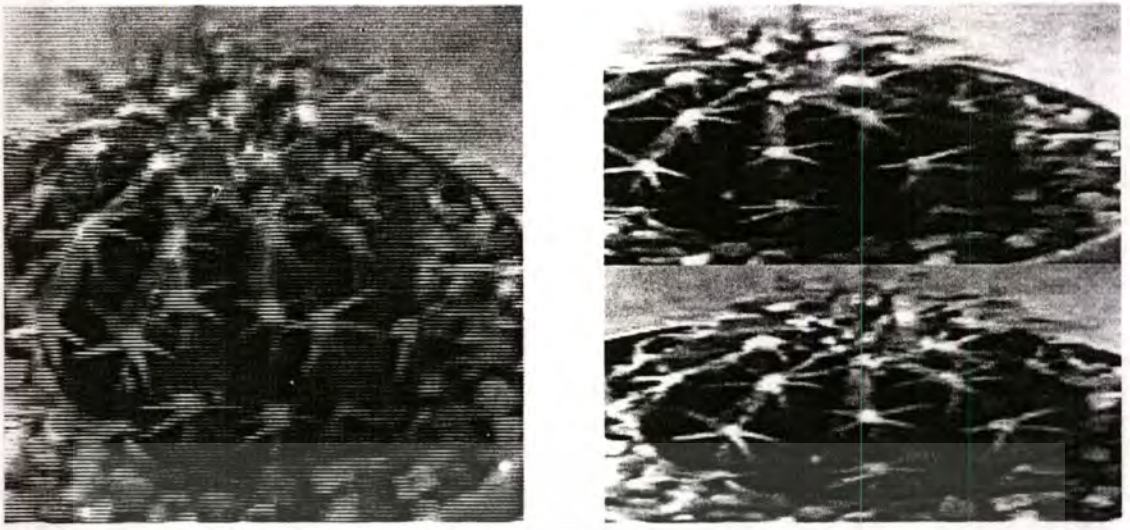
This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

on the efficient domain pool design, but such papers have emphasized iteration-free for the decoding process. B. Wohlberg [14] discussed the survey literatures related to domain pool design and concluded that the simplest way to speed up the fractal encoding is domain pool reduction. However, this method is a trade-off between domain pool size and image quality; the smaller domain pool leads to poorer image quality. Therefore, the objectives of this chapter are to design the effective domain pool whose size is reduced, as small as possible, and to retain a feasibly good image quality. For this reason, a novel approach has been proposed for designing the domain pool, i.e., mutual domain pool design. The mutual domain pools based on the *equivalent* property of two complete metric spaces are designed for speeding up the encoding and decoding time. This method provides the smaller domain pools, thus reducing the searching time of domain-range mapping. The codes generated by these mutual domain pools are called *mutual fractal codes*.

The remaining problem is that the high vertical-frequencies of the stereoscopic image pairs are suppressed. There are two methods to avoid this problem. The former is a mutual scanline fractal coding for frame-based images. In this method, the scanline technique is used to partition the range block of images. It is a simple and practical technique to preserve the high vertical-frequencies of the stereoscopic image pairs. It also avoids the image-crosstalk problem caused by blending the left and right images [36]. The latter is a mutual 2-D fractal coding (M2DFC) for field-based images. In this approach, a de-interleaved technique is utilized for eliminating the high vertical-frequencies of the stereoscopic image pairs. This technique is employed in the preprocessing step of the M2DFC and is explained in more detail in Chapter 4.

The rest of this chapter is organized as follows: Section 3.2 simply and concisely analyzes the construction of a field-sequential stereoscopic video sequence and its characteristics. The high vertical-frequencies of stereoscopic image pairs are also addressed. In Section 3.3, the theoretical model of mutual fractal coding is formed. Then the design of the coding system is proposed in Section 3.4. Section 3.5 describes the implementations and applications of the intended approach. The experimental results are illustrated in Section 3.6. Finally, the conclusion is given in Section 3.7.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



(a) (b)

Figure 3.1 (a) Original 256 x 256 pixels, 8 bpp "cactus" stereoscopic image pair in field-sequential format (frame-based image) and (b) field-based image separated from the original stereoscopic image pair.

3.2 Analysis of Field-Sequential Stereoscopic Image Sequences

In this section, we simply and concisely analyze the construction of a field-sequential stereoscopic video sequence and its characteristics.

3.2.1 Construction of Field-Sequential Stereoscopic Video Sequences

Generally, the field-sequential stereoscopic video is made up of a sequence of stereoscopic image pairs. Figure 2.3 shows its construction. This construction is compatible with the typical scanning of television systems for which every frame consists of two fields: even and odd. The number of times for repainting the screen in one second is called the *field rate*. This field rate is customarily related to a vertical synchronous frequency, 60 Hz for a NTSC system. With this field rate, it is so expeditious that the human eye cannot identify the alternation between the even and odd fields. Thus, the field-sequential scanning is an appropriate technique for occupying the left and right images in one frame, consisting of the even field for the left image and the odd field for the right image and for displaying on standard televisions and computer monitors. A model of the field-sequential stereoscopic video acquisition system is illustrated in Figure 2.4. Figure 2.5 shows a set of stereoscopic image pairs acquired by this model.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

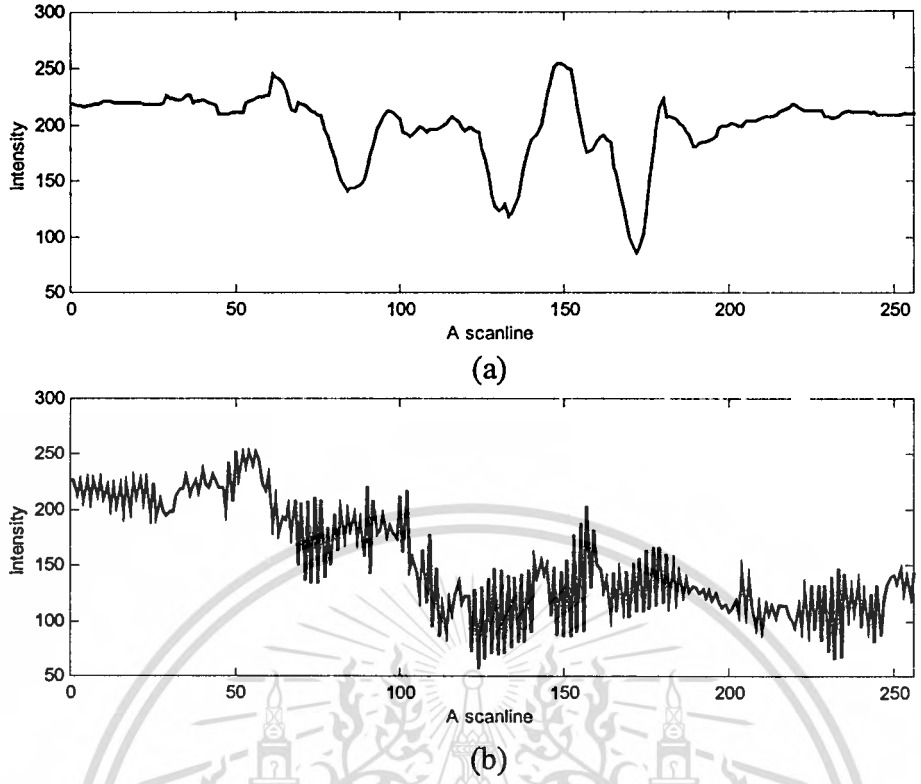


Figure 3.2 Representation of (a) the horizontal image scanline in row 51st and (b) vertical image scanline in column 120th of the original stereoscopic image pairs in Fig. 3.1 (a).

3.2.2 Characteristics of Field-Sequential Stereoscopic Image Pairs

Fundamentally, an image can be viewed as a two-dimensional signal composed of horizontal and vertical axes that denote horizontal and vertical frequencies, respectively. For a field-sequential stereoscopic image, both horizontal and vertical frequencies have a great difference as illustrated in Figure 3.2, but for a typical image, they have no great difference as graphically depicted in Figure 3.3. At this point, it is summarized that the field-sequential stereoscopic video sequence differs from the typical video sequence, in that it contains a large amount of high frequencies in the vertical direction of each frame. These high frequencies are substantial for depth perception. This characteristic of the field-sequential stereoscopic video sequence is very important for designing a coding scheme described in the following sections. However, in order to preserve the high vertical-frequencies, the MSFC can be applied to the frame-based image depicted in Figure 3.1(a) and the M2DFC introduced in Chapter 4 to the field-based image shown in Figure 3.1(b). These approaches are designed and implemented with the theoretical model described in the next section.

Forbidden to modify the content, and cite the document when use.

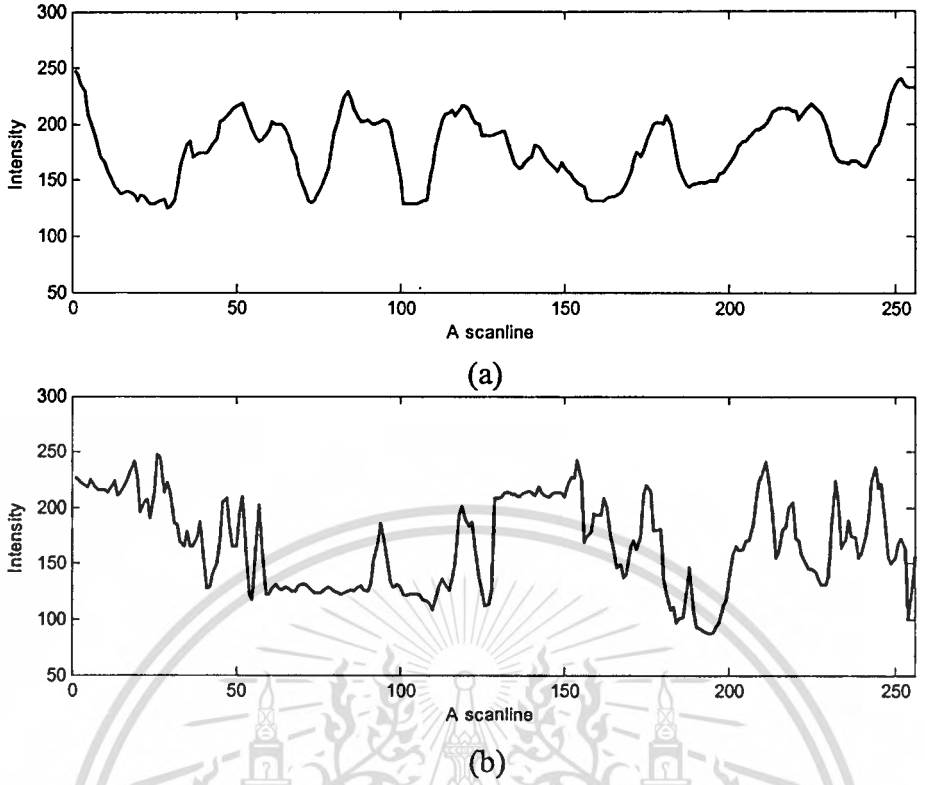


Figure 3.3 Representation of (a) the horizontal image scanline in row 51st and (b) vertical image scanline in column 60th of the field-based image in Fig. 3.1(b).

3.3 Theoretical Model

In this section, a theoretical model of mutual fractal coding is concisely presented. This model is made up of the following principal theories: metric spaces, affine transformations, contractive mapping fixed-point theorem and partitioned iterated function system (PIFS). However, we describe only the relevant definitions and theorems. Additional details of the above theories refer to [37]. In addition, only the theory of a scanline model is described. For an extended (2-D) model, the concept of design and implementation scheme is described and illustrated in Chapter 4.

3.3.1 Model of Mutual Fractal Codes

This subsection begins with describing the definition of elementary mathematical model for the mutual scanline fractal coding, that is, *complete* metric spaces of a stereoscopic image pair and their *equivalent* property. Typically, the fractal image coding requires such metric spaces to be complete in order that they

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

have no missing values. The equivalent property of two complete metric spaces is also important for designing mutual domain pools. The significant condition of the two equivalent metric spaces is that there is no extreme stretching or compression of the spaces [37]. With this condition, the stereoscopic image pairs can be viewed as a mutual deformation of images. That is, a left image of the stereoscopic image pair is a deformation of a right image and vice versa. In practice, the deformation mechanisms of a stereoscopic video acquisition system are the left and right prisms, as schematically shown in Figure 2.4. For this reason, let (X,d) be a complete metric space of the stereoscopic image pair. Then it can be decomposed into two complete metric spaces (X_L,d_L) and (X_R,d_R). The equivalence of two complete metric spaces is defined as follows:

Definition 1: *Let (X_L,d_L) and (X_R,d_R) be complete metric spaces of the left and right discrete images of a stereoscopic image pair. Then these complete metric spaces are equivalent if there is a function $f : X_L \to X_R$ that is one-to-one and onto, such that the metric $\tilde{d}_L$ on X_L defined by*

$$\tilde{d}_L(x,y) = d_R(f(x),f(y)),\ \forall x,y \in X_L$$

is equivalent to d_L. □

After introducing the definition of metric spaces with their necessary properties, the transformations of these metric spaces are addressed. In case of the mutual scanline fractal coding, the transformations used to generate the mutual fractal codes are a pair of affine transformations and are contractive mappings [37]. The condition that makes these transformations contractive is defined as follows:

Definition 2: *Let (X_L,d_L) and (X_R,d_R) be complete metric spaces and let $f : X_L \to X_R$ and $g : X_R \to X_L$ be affine transformations. These transformations f and g are called contractive mappings if there exist constants $0 \le s_L, s_R < 1$ such that*

$$d_R(f(x),f(y)) \le s_L \cdot d_L(x,y),\ \forall x,y \in X_L$$

and

$$d_L(g(x),g(y)) \le s_R \cdot d_R(x,y),\ \forall x,y \in X_R\ .$$

Such constants s_L and s_R are called contractivity *factors for f and g, respectively.* □

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Theorem 1: *From Definition 2, the contractivity factor of* $(f \circ g)(x) = f(g(x))$, *where* $f \circ g$ *is the composition of* f *and* g, *is* $s_L \cdot s_R$.

Proof: By Definition 2, we have

$$d_R(f(x), f(y)) \le s_L \cdot d_L(x, y), \ \forall x, y \in X_L \tag{3.1}$$

and

$$d_L(g(x), g(y)) \le s_R \cdot d_R(x, y), \ \forall x, y \in X_R, \tag{3.2}$$

where s_L and s_R are contractivity factors for f and g, respectively. We begin with inequality (3.2) and multiply both sides by contractivity factor s_L. The result becomes

$$s_L \cdot d_L(g(x), g(y)) \le s_L \cdot s_R \cdot d_R(x, y). \tag{3.3}$$

However, we also have

$$d_R(f(g(x)), f(g(y))) \le s_L \cdot d_L(g(x), g(y)) \le s_L \cdot s_R \cdot d_R(x, y). \tag{3.4}$$

By the definition of contractive mapping, $s_L \cdot s_R$ is contractivity of $f \circ g$. This proof is now complete. □

Definitions 1 and 2 provide the suitable complete metric spaces and contractive transformations, respectively, for generating mutual fractal codes. Based on these definitions, a pair of contractive transformations f and g is imposed to mutually map between two complete metric spaces (X_L, d_L) and (X_R, d_R). Theorem 1 also proves that $(f \circ g)(x)$ is contractive and its contractivity factor is $s_L s_R$. However, in order to solve the *image encoding problem* (or *inverse problem*) [14], [34], the fixed-point theorem is very important. Based on Definitions 1 and 2, and Theorem 1, the contractive mapping fixed-point theorem presented in [14], [34] can be extended to the objective as Theorem 2.

Theorem 2: *Let* (X_L, d_L) *and* (X_R, d_R) *be equivalent complete metric spaces, and let* $f : X_L \to X_R$ *and* $g : X_R \to X_L$ *be contractive mappings. Then* $f \circ g$ *possesses exactly one fixed point* $x_f \in X$, *such that* $X = X_L \cup X_R$ *and* $X_L \cap X_R = \varnothing$. *Furthermore, for any point* $x \in X$, *the sequence*

$$\left\{ (f \circ g)^{\circ n}(x) : n = 0, 1, 2, \ldots \right\}$$

converges to x_f. *That is,* $x_f = (f \circ g)(x_f) = \lim_{n \to \infty} (f \circ g)^{\circ n}(x)$, *for each* $x \in X$.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Proof: Let $X=X_L\cup X_R$ and $X_L\cap X_R=\varnothing$. Then, based on Definition 2, f and g are contractive transformations that mutually transform between two equivalent complete metric spaces (X_L,d_L) and (X_R,d_R). To prove that the sequence

$$(f\circ g)^{\circ n}(x)=\underbrace{f(g(f(g(\cdots(x)\cdots))))}_{n \text{ times}} \tag{3.5}$$

converges to a fixed point. By Theorem 1, let $s=s_L\cdot s_R$ and let $p>q$. Then for $x\in X$,

$$\begin{aligned} d\left((f\circ g)^{\circ q}(x),(f\circ g)^{\circ p}(x)\right) &< s\cdot d\left((f\circ g)^{\circ(q-1)}(x),(f\circ g)^{\circ(p-1)}(x)\right) \\ &< s^q\cdot d\left(x,(f\circ g)^{\circ(p-q)}(x)\right). \end{aligned} \tag{3.6}$$

Using triangle inequality, the following is obtained

$$\begin{aligned} d\left(x,(f\circ g)^{\circ(p-q)}(x)\right) &\le d\left(x,(f\circ g)^{\circ(p-q-1)}(x)\right)+d\left((f\circ g)^{\circ(p-q-1)}(x),(f\circ g)^{\circ(p-q)}(x)\right) \\ &\le d(x,(f\circ g)(x))+d\left((f\circ g)(x),(f\circ g)^{\circ 2}(x)\right) \\ &\quad +\cdots+d\left((f\circ g)^{\circ(p-q-1)}(x),(f\circ g)^{\circ(p-q)}(x)\right) \\ &\le \left(1+s+\cdots+s^{p-q-1}\right)d(x,(f\circ g)(x)) \\ &\le \frac{1}{1-s}d(x,(f\circ g)(x)). \end{aligned} \tag{3.7}$$

Substituting (3.7) to (3.6), the following is obtained

$$d\left((f\circ g)^{\circ q}(x),(f\circ g)^{\circ p}(x)\right)<\frac{s^q}{1-s}d(x,(f\circ g)(x)). \tag{3.8}$$

From Eq.(3.8), if p and q are sufficiently large and $s<1$, the left hand side of Eq.(3.8) is made as small as possible. This implies that the sequence

$$\{x,f(g(x)),f(g(f(g(x)))),\ldots\}$$

is a Cauchy sequence. Thus, the limit point

$$x_f=\lim_{n\to\infty}(f\circ g)^{\circ n}(x) \tag{3.9}$$

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

is in X, such that X is complete. Since $f \circ g$ is contractive as described in Theorem 1, it is continuous, and hence

$$(f \circ g)(x_f) = (f \circ g)\left(\lim_{n\to\infty}(f \circ g)^{\circ n}(x)\right) = \lim_{n\to\infty}(f \circ g)^{\circ(n+1)}(x) = x_f . \qquad (3.10)$$

Moreover, let x_{f1} and x_{f2} be two fixed points of $f \circ g$. Then

$$x_{f1} = (f \circ g)(x_{f1}),\ x_{f2} = (f \circ g)(x_{f2})$$

and

$$d(x_{f1}, x_{f2}) = d((f \circ g)(x_{f1}), (f \circ g)(x_{f2}))$$

$$\le s\,d(x_{f1}, x_{f2})$$

$$(1-s)d(x_{f1}, x_{f2}) \le 0, \qquad (3.11)$$

which implies $d(x_{f1}, x_{f2}) = 0$, hence $x_{f1} = x_{f2}$. The proof is now complete. □

Theorem 2 proves that $f \circ g$ still possesses exactly one fixed point, even though it performs on two equivalent complete metric spaces. In practice, this provides an effective design of mutual domain pools in Subsection 3.4.2. Now, Theorem 2 is applied to a *partitioned iteration function system* (PIFS). In this chapter, we also make use of divide-and-conquer technique to provide suitable conditions for the PIFS. This technique involves solving a particular computational problem by *dividing* it into more subproblems of smaller size and then *merging* the solutions of the subproblems to produce a solution of the original problem. For scanline case, we first divide an image into a number of pairs of scanlines. Then, each pair of scanlines is defined as follows: D_L and D_R denote the mutual domain pools of scanlines of left and right images, respectively, whereas R_L and R_R denote the scanlines of range blocks. Finally, the following definition generates the mutual fractal codes.

Definition 3: *Let (X_L, d_L) and (X_R, d_R) be complete metric spaces of the left and right discrete images of a stereoscopic image pair, and let $D_L \subset X_L$ and $D_R \subset X_R$. Then a* partitioned iteration function system (PIFS) *is a collection of contractive mappings:*

$$w_u : D_{L(u)} \to R_R \ \text{for } u = 1,2,\ldots,N$$

and

$$w_v : D_{R(v)} \to R_L \ \text{for } v = 1,2,\ldots,M,$$

where N and M are a number of domain blocks of D_L and D_R, respectively. □

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Definition 3 makes fractal image encoding practical. Therefore, the mutual fractal codes for a whole image can be generated by means of the following definition.

Definition 4: *By Definition 3, the mutual fractal codes W of an encoded image can be obtained as*

$$W = \bigcup_{k=1}^{K/2}\left(\left(\bigcup_{u=1}^{N} w_{u,k}\right)\cup\left(\bigcup_{v=1}^{M} w_{v,k+1}\right)\right),$$

where N and M are a number of range blocks of R_L and R_R, respectively, and K is a total scanline of a stereoscopic image pair. □

Now, the theory can be made more practical for generating mutual fractal codes based on the previous definitions and theorems. The following subsections describe the principle of mutual scanline fractal encoding and decoding.

3.3.2 Mutual Scanline Fractal Encoding

Encoding is a step to encode the mutual fractal codes generated by Definitions 1 to 4 and Theorem 2. The heart of this phase is to determine the optimal values of parameters of domain-range mappings. To account for manipulation of the mutual scanline fractal encoding, there are two issues that are introduced. First, we assume that there are mutual domain pools D_L and D_R, which denote the scanlines of a stereoscopic image pair, and their intersection ($D_L \cap D_R = \varnothing$) is empty. This implies that the mutual domain pools are not unified in spatial coordinates. The mutual domain pools satisfy Definition 1. Second, there is a pair of affine transformations $f : D_L \rightarrow R_R$ and $g : D_R \rightarrow R_L$, where R_R and R_L are the scanlines of range blocks and these transformations satisfy Definition 2.

In practice, the process of generating mutual fractal codes based on Definition 3 requires two types of data blocks, namely, domain and range blocks. Usually the size of the domain block is greater than that of the range block. For establishing range blocks, each scanline of the original image is partitioned into a set of non-overlapping blocks. For setting up mutual domain pools, where the domain blocks belong, each scanline of the original image is partitioned into a set of overlapping blocks, thus, the

This material is reserved for educational use only. Not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

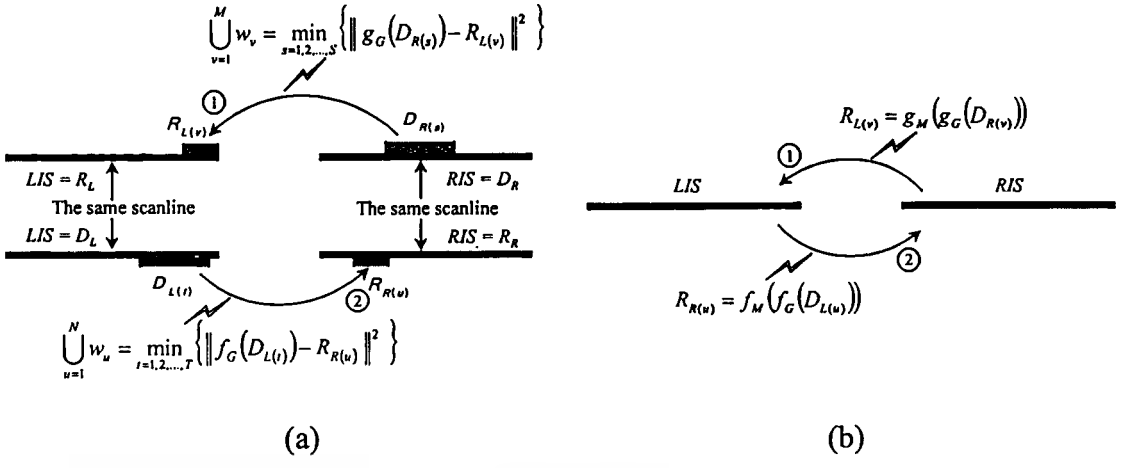


Figure 3.4 (a) Encoding and (b) decoding steps of mutual fractal codes.

mutual domain pools are a collection of overlapping blocks. After providing the appropriate domain and range blocks, domain-range-mapping process is reached. In this process, there are two main steps for mutually obtaining the best matching between the domain and range blocks. As depicted in Figure 3.4(a), the mutual domain pools consist of D_L and D_R. Each range block of R_L and R_R mutually searches for the best matching of a domain block in the mutual domain pools D_R and D_L, respectively. That is, in the first step, each range block of the scanline $R_{L(v)}$ searches for the best matching of the domain blocks in the domain pool $D_{R(s)}$ by means of

$$\bigcup_{v=1}^{M} w_v = \min_{s=1,2,\ldots,S} \left\{ \left\| g_G\left(D_{R(s)}\right) - R_{L(v)} \right\|^2 \right\}, \tag{3.12}$$

where S is a total possible domain block in the domain pool D_R, and $g_G(.)$ denotes the geometric transformation described in the end of this subsection. In the second step, each range block $R_{R(u)}$ searches for the best matching of a domain block in the domain pool $D_{L(t)}$ by means of

$$\bigcup_{u=1}^{N} w_u = \min_{t=1,2,\ldots,T} \left\{ \left\| f_G\left(D_{L(t)}\right) - R_{R(u)} \right\|^2 \right\}, \tag{3.13}$$

where T is a total possible domain block in the domain pool D_L, and $f_G(.)$ denotes the geometric transformation.

The tool used to measure the similarity of the domain and range blocks is composed of geometric and massic transformations. In Eqs.(1) and (2), geometric transformations are $g_G(.)$ and $f_G(.)$ whereas massic transformations are implicit in the form of least-square error (Subsection 3.4.4 describes more details in this issue.) The

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

step of generating mutual fractal codes W is that the domain blocks are transformed by using contractive affine (geometric) transformations. Then the root-mean-square (rms) metric (massic transformation) is applied to measure the similarity of the domain and range blocks. Finally, the mutual fractal codes generated by these transformations are uniformly quantized so as to be compact forms.

3.3.3 Mutual Scanline Fractal Decoding

The concept of decoding mutual fractal codes is very simple. It involves the obtaining of a fixed point of a contractive transformation. That is, the decoding process is an iteration of the contractive transformation on any initial values until it reaches an approximate fixed point. However, the mutual scanline fractal decoding differs from the conventional one. In addition, the concept of decoding is based on the fact that if the seed image is initialized close to the original image, then the contractive transformation converges rapidly to a fixed point [38]. In order to describe the manipulation of this decoding process, the left image scanline (*LIS*) and right image scanline (*RIS*) are set up with zero-initial values, i.e., all pixels of the scanlines are zeroes. The seed image scanlines are *LIS* and *RIS*. In each iteration, there are two steps to decode the mutual fractal codes as depicted in Figure 3.4(b). In the first step, all range blocks of the *LIS* are mapped with the appropriate domain-transformed blocks from the *RIS*, such that $R_{L(v)}=g_M(g_G(D_{R(v)}))$, where $g_G(.)$ and $g_M(.)$ are contractive (geometric) and massic transformations and $D_{R(v)}$ is a domain block of the *RIS*. In the second step, the outcome in the *LIS*, which is a domain-block scanline of the *RIS*, is performed with the same process as the first step. Note that the *LIS*, a seed image scanline of *RIS*, is not initialized with zero values, but holds the results from the first step. This implies that the seed image scanline *LIS* is initialized close to the original one. Thus, the transformation $g_M(f_G(.))$ converges rapidly to a fixed point. This is an effective method for generating mutual fractal codes that can speed up decoding time.

3.4 Design of Coding System

In the designing phase, there are five significant issues that have to be considered and decided for a coding scheme. These issues are scanline partition, mutual domain pool design, mutual scanline transformations, distortion measure and

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

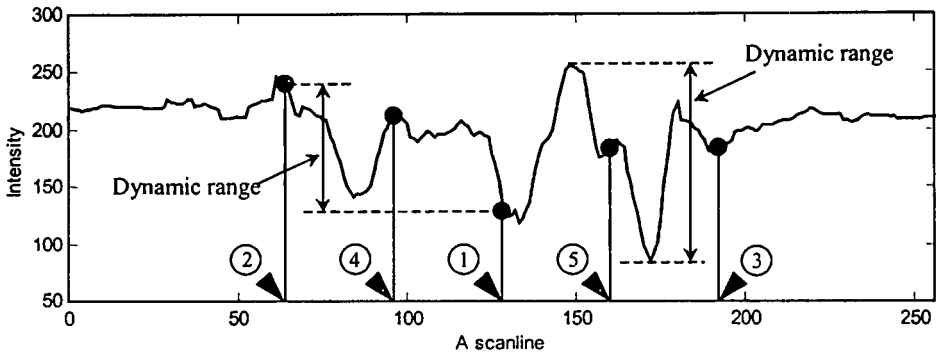


Figure 3.5 Illustration of adaptive-size partition by using dynamic range, which refers to the extremes of intensity change of an image scanline in any intervals.

bit allocation. All the issues described in this section can be extended easily to the 2-D case described in Chapter 4.

3.4.1 Range Block Partition: Scanline Partition

Scanline partition is a procedure to divide the scanline of the original image into non-overlapping one-dimensional (1-D) blocks. These blocks are called *range* blocks. The size of the range block is an important factor that directly influences bit rate. This means that the larger range blocks lead to a lower bit rate. On the other hand, the smaller size captures more details of image scanlines.

There are two kinds of partitions: fixed- and adaptive-size partitions. For the first kind of partition, each scanline is equally divided into a fixed size, such as 1x8, 1x16, or 1x32. This method is very simple, but not flexible for the bit rate. For the second kind of partition, the binary-tree partition technique is used to divide the image scanline. In this method, each scanline is equally divided into two subparts. Then the *dynamic* range, which refers to the extremes of intensity change of an image scanline in any interval as depicted in Figure 3.5, is used to decide splitting. If the dynamic range value of each subpart exceeds the threshold, then that subpart is equally divided, otherwise, it does nothing. This task is repeatedly processed in the previous fashion until reaching the restricted block size. In Figure 3.5, the numbers in the circle indicate the sequence of partition. Hence, the dynamic range threshold is an important parameter to control a number of range blocks. For the lower bit rate and image quality, the smallest and largest sizes of bounded range blocks are defined, i.e., 1x8

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

and 1x64 are of the smallest size and the largest size, respectively. By this method, the adaptive partition makes the bit rate flexible and also achieves better image quality.

3.4.2 Mutual Domain Pool Design: A One-dimensional Approach

A domain pool is a place in which the range block searches for the best matching block. A good design of the domain pool leads to less encoding time of the fractal coding [14]. In this chapter, a simple method is proposed for designing domain pools by utilizing the equivalent property of two complete metric spaces, as described by Definition 1 in Subsection 3.3.1. The coding scheme is designed with two domain pools, D_L and D_R. These domain pools are independent, but they are equivalent. In the searching step, each range block on the space R_L finds the best matching domain block in the space D_R and vice versa. This technique makes the domain pool size smaller, thus reducing the encoding time.

3.4.3 Geometric Affine Transformation: Mutual Scanline Transformations

In this subsection, a pair of scanline transformations designed for our codec (coder and decoder) is proposed. These transformations are 1-D affine transforms defined as:

$$\begin{cases} f_G(x_l) = ax_l + b, & \forall x_l \in X_L \\ g_G(x_r) = cx_r + d, & \forall x_r \in X_R \end{cases} \tag{3.14}$$

where a and c are geometrical contractivity factors [7], [11], [13] of the transformations, b and d are translation factors, and x_l and x_r are range blocks on the complete metric spaces (X_L,d_L) and (X_R,d_R), respectively. In practice, a and c are set to a constant value, usually equal to 0.5, and b and d are integers. $f_G(x_l)$ and $g_G(x_r)$ are domain-transformed blocks on the complete metric spaces (X_L,d_L) and (X_R,d_R), respectively. The steps of generating mutual fractal codes are as follows. First, each domain block whose size is greater than the range block in the metric space (X_L,d_L) is transformed by f_G in order to find the best matching of the range block in the metric space (X_R,d_R). At the same time, each domain block in the metric space (X_R,d_R) is transformed by g_G in order to find the best matching of the range block in the metric space (X_L,d_L). In this case, the tool for measuring the best matching block is root-mean-square (rms) metric presented in the next subsection.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

3.4.4 Intensity Affine Transformation: Distortion Measure

Distortion measure is a tool to measure the similarity of two image scanlines. In practice, the tool commonly used for measuring the distortion of domain-range mapping is the root-mean-square (rms) metric. This metric is easy to compute the optimal values of parameters s_j and o_j, where s_j and o_j denote contrast and brightness, respectively, in Eq.(4) [39]. Let d_i and r_i be elements of 1-D domain and range blocks, respectively, we can obtain the minimal value rms_j that occurs when s_j and o_j are optimized by means of

$$rms_j = \left[\sum_i \left(s_j \cdot d_i + o_j - r_i \right)^2 \right]^{\frac{1}{2}}, \tag{3.15}$$

where i and j are pixel's order within a range block and the range block's order within a scanline, respectively. To determine the optimal values of s_j and o_j, the method of *least squares* is used. That is, the partial derivatives of square of Eq.(3.15) with respect to s_j and o_j are calculated, respectively, and then are set to zero, so that the optimal values s_j and o_j are computed by using two linear equations. As a result, the optimal equations of parameters s_j and o_j are obtained as follows:

$$s_j = \frac{n \sum_i d_i r_i - \sum_i d_i \sum_i r_i}{n \sum_i d_i^2 - \left(\sum_i d_i \right)^2} \tag{3.16}$$

and

$$o_j = \frac{1}{n} \left(\sum_i r_i - s_j \sum_i d_i \right), \tag{3.17}$$

where n is a number of pixels in a range block. However, if the denominator of Eq. (3.16) is zero, then we define $s_j = 0$ and $o_j = \frac{1}{n} \sum r_i$. In practice, the contractivity factor s_j of a massic class is sensitive to the convergence condition. To guarantee the contractivity criterion, s_j is limited by a maximum value, i.e., $|s_j| \leq S_{max}$, where S_{max} is a maximum boundary of s_j. The details of convergence condition can be found in [40].

3.4.5 Bit Allocation

Bit allocation is a process to appropriately define a number of bits for each encoded parameter, so that the reconstructed image is close to the original one. In this

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

chapter, the following parameters, domain-block location (D_l), domain-block orientation (D_o), range-block size (R_s), scaling (s), and offset (o), represent the mutual fractal codes generated from a pair of contractive affine transformations and massic transformations. These parameters can be classified into two classes: geometric and massic classes according to their transformations [34], [41]. The geometric class is the outcome from a scanline partition and transformation. Its bit allocation depends on the geometrical characteristic of images, such as image size, range-block size and others, therefore, it is almost restricted. On the other hand, the massic class is the result of measuring similarity or distortion of domain-range mapping by means of root-mean-square metric. Its bit allocation is very flexible and affects directly to the image quality, however, fewer bits for encoded parameters (massic class) trade with the poorer image quality. For a suitable design, the bit requirement of mutual fractal codes is allocated as follows: For geometric class, domain-block location (D_l), domain-block orientation (D_o), and range-block size (R_s) are required 8, 1, 2 bits, respectively. For massic class, the scaling (s) and offset (o) parameters based on the optimal values are presented in [39]. These parameters require 5 and 7 bits, respectively. The total bit allocation for each range block can be obtained as follows:

$$B_{MSFC} = D_l + D_o + R_s + s + o \tag{3.18}$$

Note that in case of fixed-size partition, the factor R_s is disregarded, since all range-block sizes are equal, therefore the bit allocation for each range block is reduced by this factor.

3.4.6 Complexity Reduction of Mutual Scanline Domain Pools

In this subsection, the complexity of searching algorithms for domain-range mapping is first analyzed. All algorithms implemented in this section are based on a first-match searching technique. Here the domain pool size (length) is N_D. If the first target (domain block) in the list of a domain pool is matched, then the number of comparisons is 1. This is the best time for a successful search. On the other hand, if the last target in the list is matched, then the number of comparisons is N_D. This is the worst case for a successful search, therefore, in case of average successful search, the algorithm complexity can be modeled as

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

$$\frac{1}{N_D}\sum_{i=1}^{N_D} i = \frac{1+2+3+\cdots+N_D}{N_D} = \frac{N_D+1}{2}. \tag{3.19}$$

At this point, It can be summarized that the complexity of the algorithm based on first-match searching is $O(N_D)$ [42]. It is evident that the first-match searching algorithm has the same complexity as the exhaustive searching algorithm, whose complexity is also $O(N_D)$. These algorithms are difference in that the exhaustive search-ing always reaches N_D times of comparisons, whereas the first-match searching is possibly less than N_D times of comparisons. Thus, the first-match searching ordinarily runs faster than the exhaustive searching. In case of the domain pool design, the domain pools of images are dependent on their characteristics. That is, the different images yield different domain pools. This characteristic makes the domain pools adaptable. For this reason, the binary searching technique whose complexity is $O(\log_2 n)$ [43] is not suitable for direct application to the domain-range mapping procedure, since the domain pools have to be sorted every time. In order to avoid these obstacles, the reduction of the domain pool size is considered; the smaller domain pool leads to lower complexity. In this case, to verify the efficiency of the domain pool design between the conventional and mutual-scanline domain pools, the domain pool sizes are ordinarily obtained by

$$N_D = (M-2R)(N-2R), \tag{3.20}$$

where M and N represent horizontal and vertical image dimensions, respectively, and R denotes the range block size. In the case of mutual scanline domain pools, the vertical range-block dimension is set to 1 and the searching space satisfies Definition 3. The number of domain blocks N_D is reduced to M-2R blocks. When compared to the conventional one, the mutual scanline domain pool is N-2R times smaller. In the worse case for a successful search, the comparison is M-2R times. This implies that the domain-range-mapping procedure can speed up more than M-2R times. Table 3.1 shows the numerical comparison based on the 256x256 image size.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Table 3.1 Comparison of domain pool sizes of the conventional and proposed methods based on 256x256 image size.

Method	Range block size (R)	Domain pool sizes (N_D)
Proposed method	1x4	248 (1x8)
	1x8	240 (1x16)
	1x16	224 (1x32)
Conventional method	4x4	61,504 (8x8)
	8x8	57,600 (16x16)
	16x16	50,176 (32x32)

3.5 Implementation and Application to Field-Sequential Stereoscopic Image Pairs

In this section, we have introduced the mutual domain pool design one-dimensional (scanline) implementation scheme of mutual fractal coding and its application to field-sequential stereoscopic image sequences. The implementation scheme of mutual scanline fractal coding is illustrated in Figure 3.6. First the manipulation of the implementation scheme with the encoder section is described. In this scheme, a divide-and-conquer technique is used. The reason for applying this technique is given in Subsection 3.3. The dividing procedure is a preprocessing step that splits a whole image into a number of pairs of scanlines, then, each pair of scanlines is divided by a scanline partition. This scanline partition is an important procedure, since it defines the range-block size that is a significant factor to control bit rate and image quality. In the case of an adaptive-size partition of range blocks, the dynamic range threshold is a parameter used to control a number of range blocks. In other words, we can also control the bit rate of this scheme by using this threshold. The outcome of this procedure is composed of two sets of non-overlapping range blocks, i.e., R_L and R_R, and two sets of overlapping domain blocks, i.e., D_L and D_R. In this case, there are two domain pools (D_L- and D_R-domain pools) for keeping the sets of overlapping domain blocks and these domain pools are called *mutual domain pools*. After that all range blocks of R_L and R_R search for the matching blocks in the domain pool D_R and D_L by two domain-range-mapping algorithms $f : D_L \rightarrow R_R$ and $g : D_R \rightarrow R_L$, respectively. The domain-range-mapping algorithm consists of three components — geometric transformation, massic transformation and bit allocation — for

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

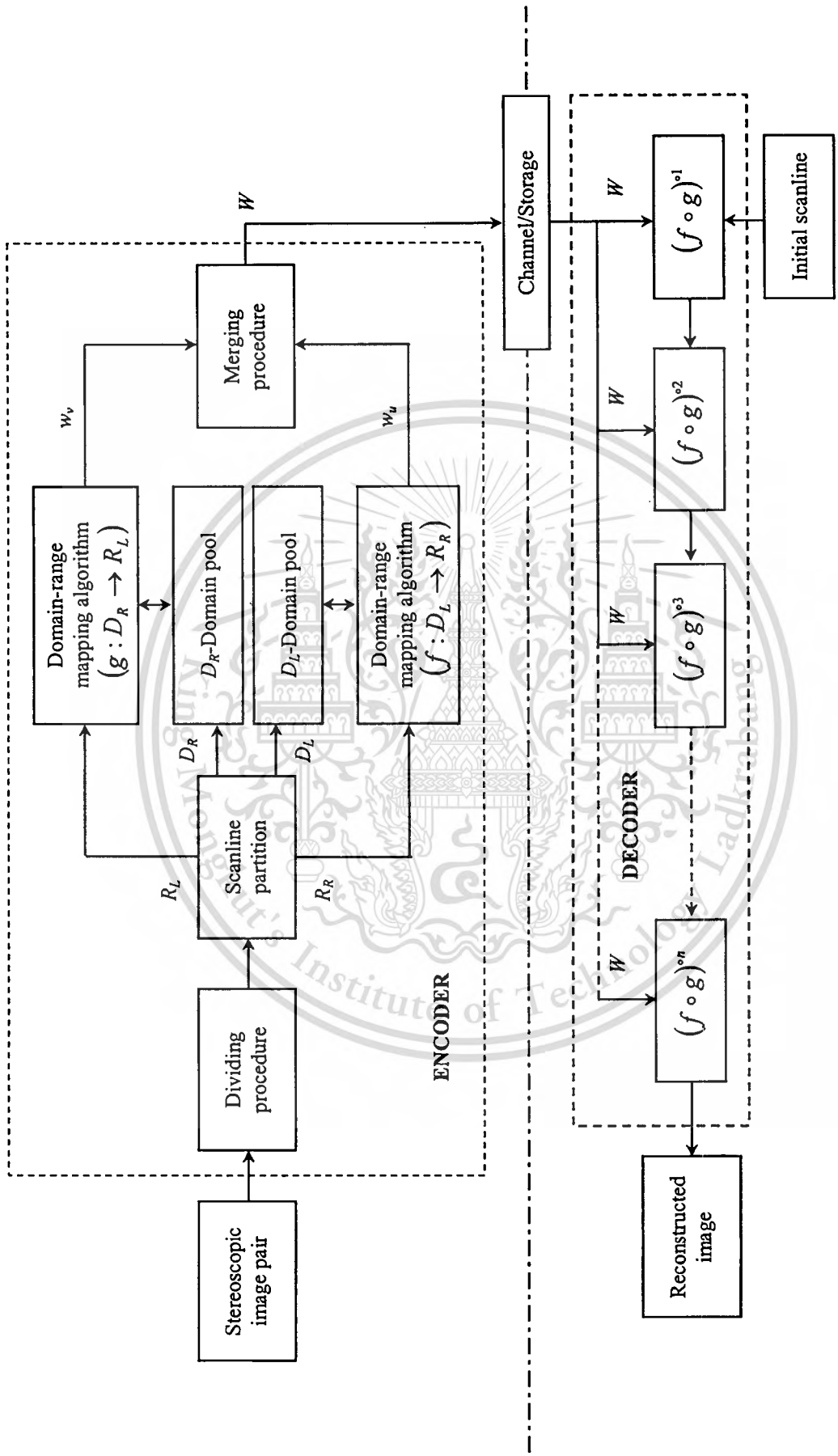


Figure 3.6 Overall implementation scheme of mutual scanline fractal coding.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

performing and generating mutual fractal codes. For a clear explanation, the relationship of geometric and massic transformations is constructed. That is,

$$\begin{cases} f = (f_M \circ f_G)(x) = f_M(f_G(x)) \\ g = (g_M \circ g_G)(x) = g_M(g_G(x)) \end{cases} \tag{3.21}$$

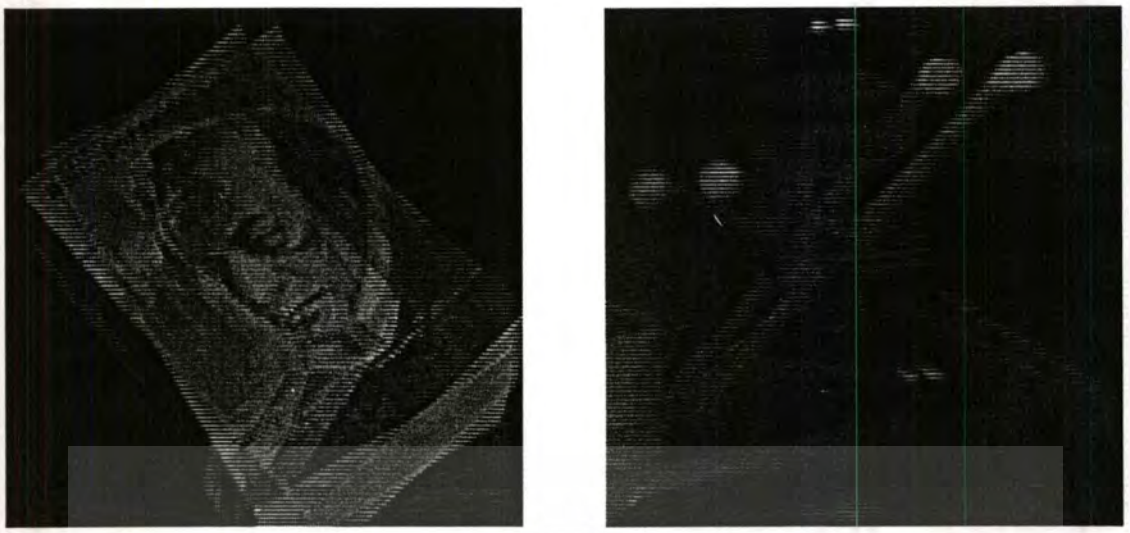
where subscripts M and G denote the massic and geometric transformations, respectively. From Eq.(3.19), f_G and g_G are 1-D affine transformations as described in Subsection 3.4.1. To obtain the best matching of domain-range mapping, all domain block sizes are shrunk to range block sizes by pixel averaging, then each range block is matched with the shrunken domain block. In this step, the massic transformation is applied to measure the similarity of those domain and range blocks. In this case, the root-mean-square metric (Eq.(3.15)) is used and its minimum value is decided as to which pair is the best matching of the domain-range mapping, then, the scaling s and offset o are retained as a part of the mutual fractal codes. However, since s and o computed by Eqs.(3.16)-(3.17) are real and are infinite numbers, they are quantized to be finite numbers. In this chapter, the uniform quantization is applied and a suitable resolution is defined as given in Subsection 3.4.5. Finally, the merging procedure integrates the mutual fractal codes w_u and w_v from both domain-range-mapping algorithms and tabulates them as $W=(D_l, D_o, R_s, s, o)$, where these symbols denote domain-block location, domain-block orientation, range-block size, scaling and offset, respectively.

For the decoder section, the operation of each block inside the decoder can be described by Figure 3.4. After decoding mutual fractal codes with a few iterations, as depicted in Figure 3.6, the reconstructed image appears.

3.6 Experimental Results

In this experiment, the performance of the intended codec is tested and evaluated in following issues: (i) encoding and decoding time and (ii) image quality. The condition of comparisons of the codec performance is based on the fractal coding on a spatial domain. The mutual scanline fractal coding (MSFC) and conventional fractal coding algorithms are implemented and simulated by using MATLAB programming. The machine for performing the simulation program is the desktop PC with PENTIUM 4 processor and the clock speed is 2.4 GHz.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



(a) (b)

Figure 3.7 Original test images: (a) a "Book" image and (b) a "Chopstick" image.

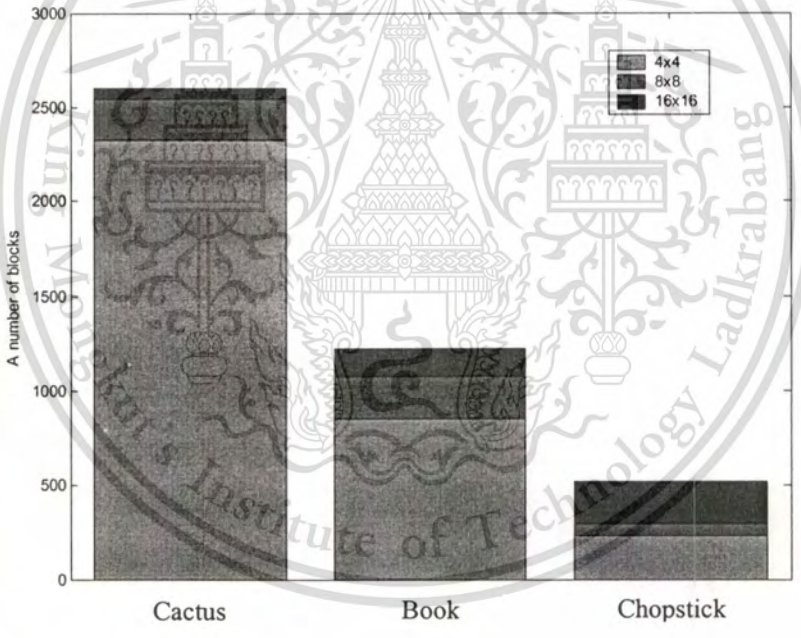


Figure 3.8 Image-complexity comparison of the test images.

In order to testify and evaluate the performance of the intended and conventional codecs, we have prepared a set of test images, i.e., "Cactus", "Book", and "Chopstick" images. The image quality is evaluated by means of peak signal-to-noise ratio (PSNR) defined as

$$\mathrm{PSNR} = 20\log\left(\frac{255}{\frac{1}{M\times N}(x-\hat{x})^2}\right) \tag{3.22}$$

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

where x and $\hat{x}$ are the original and reconstructed images, respectively, and M and N are image dimensions. For coding efficiency, we evaluate it in the form of bit rate B_R that can be calculated by the following equation:

$$B_R = \frac{\text{Total size in bits of encoder output}}{\text{Total size in pixels of encoder input}} \tag{3.23}$$

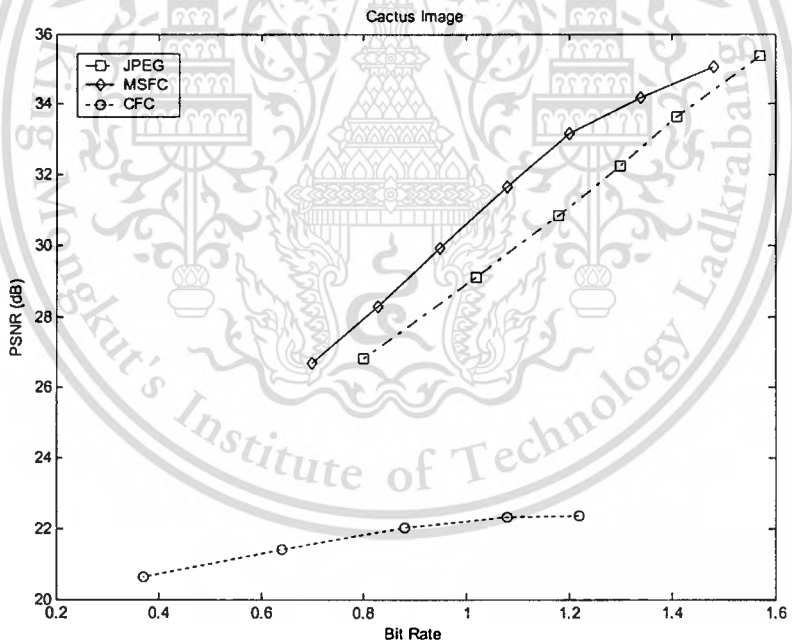
In case of the MSFC, the *total size in bits of encoder output* is the product of the total range blocks and their bit allocation. Hence, the bit allocation for each range block can be obtained by means of Eq.(3.18) for the MSFC.

Prior to experiment, the complexity of the test images is analyzed and summarized in Figure 3.8. In this case, we use the variance (dynamic range threshold) of images to classify the complexity regions into three levels, 4x4-, 8x8-, and 16x16-blocks, respectively. These different square blocks, which imply that smaller blocks, mean more complicated regions. On the other hand, larger blocks mean more smooth regions. In Figure 3.8, we can summarize that the "cactus" image is the highest complexity image, since it contains smaller square blocks (4x4 blocks) when compared to the other two images. The "book" and "chopstick" images are the median and low complexity images, respectively.

In the first experiment, we have evaluated the performance in the form of frame-based coding of the MSFC codec compared with the conventional fractal coding (CFC) and the JPEG standard coding codecs. In this experiment, the specifications of the MSFC and CFC codecs are set as follows. For the MSFC codec, the scanline technique is utilized to partition the range blocks with three-block-size adaptive partition (1x8, 1x16, and 1x32 block sizes). The first-match searching technique for domain-range mapping is selected. The bit requirement for each mutual fractal code is totally 23 bits as defined in Subsection 3.4.5. For the CFC codec, the basic and classical approach [9] with quadtree (three-level block sizes: 4x4, 8x8, and 16x16-blocks) partition is used. The image itself is assigned to be a domain pool. The searching technique is the same as the one that the MSFC is utilized. The bit allocation requires 33 bits for each fractal code. Then, the MSFC and CFC are applied to a set of test images, which are frame-based images. The fractal codes generated from both codecs are encoded by Huffman coding for the reduction of their

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

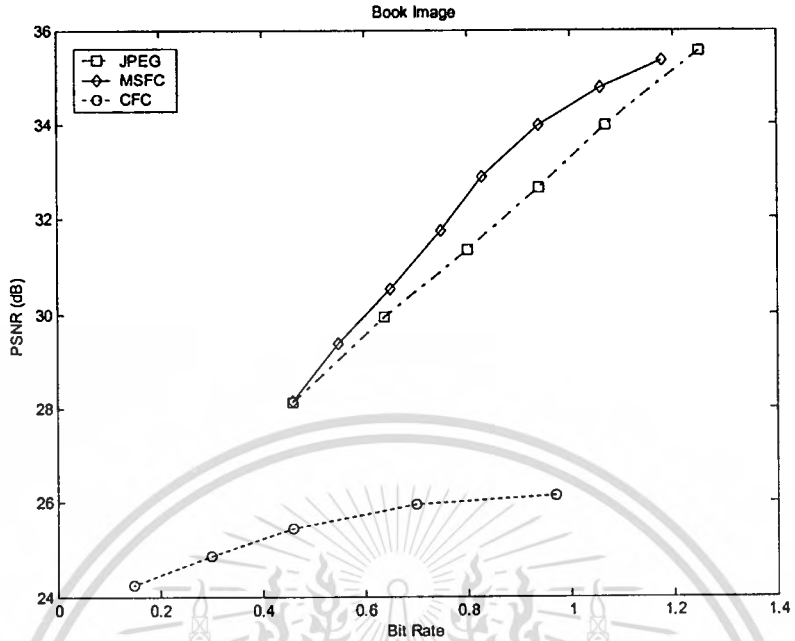
redundancies and for the fairness of comparison with JPEG standard. The experimental results graphically illustrated in Figure 3.9 indicate that in case of high and median complex images, i.e., "Cactus" and "Book" images, the MSFC codec is superior to both CFC and JPEG-standard codecs. For low complex image, the MSFC outperforms the JPEG for the bit rate greater than 0.45 bpp (See Figure 3.9(c)). This is because discrete cosine transform (DCT) carries out the better energy compaction for low complex images. In this case, for very low bit rate, only the DC coefficient is kept, thus reducing more bit rate. Simultaneously, this easily leads to the evident image crosstalk. However, in case of encoding and decoding time, the MSFC is greatly improved when compared with the CFC codec as shown in Figure 3.10 (See also Table 3.2).



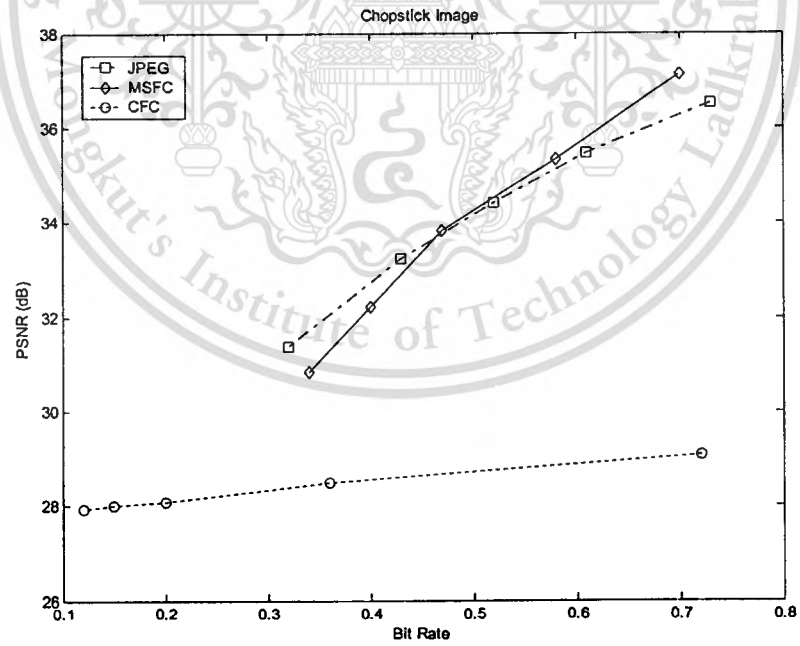
(a)

Figure 3.9 Comparison of the frame-based reconstructed images of JPEG, CFC with10 iterations, and MSFC with 5 iterations.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



(b)



(c)

Figure 3.9 (Cont.)

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

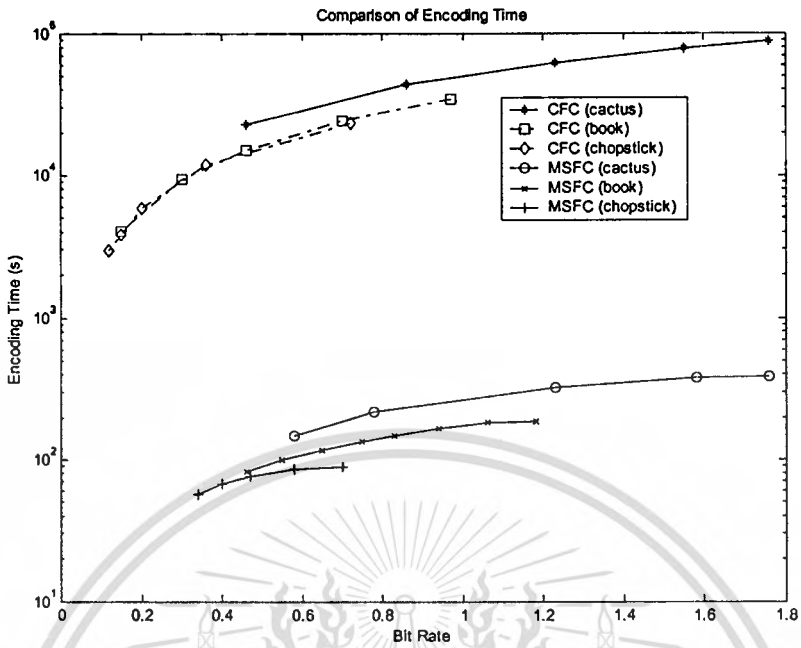


Figure 3.10 Comparison of encoding time of CFC and MSFC in frame-based coding.

Table 3.2 Comparison of average encoding time of MSFC and CFC.

Test image	MSFC (s)	CFC (s)	Improved Ratio
Cactus	290.50	58,523.76	201
Book	139.82	17,388.55	124
Chopstick	74.94	9,552.05	127

3.7 Conclusion

In this chapter, we have proposed the effective domain pool design, which is an important phase to speed up the encoding and decoding time of fractal coding. This design is called *mutual domain pool design* based on the equivalent property of two complete metric spaces. The mutual domain pool design leads to the effective coding system which can be summarized as follows: (i) The domain-pool size is reduced, thus decreasing the encoding time. (ii) Saving the number of bits for storing the

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

domain-block position, only the column coordinate of domain-pool position is kept, whereas the row coordinate is omitted, since it implicitly is with the range-block position. (iii) The mutual fractal codes generated from the mutual domain-range mapping provide faster decoding of the reconstructed images. Furthermore, to evidently evaluate the performance of the mutual domain pools without any bias, all schemes, including CFC codec, use the simple searching strategy, i.e., the first-match searching technique, for domain-range mapping. The experimental results show that in case of frame-based images the MSFC is superior to the JPEG-standard and CFC codecs in terms of image quality as illustrated in Figure 3.9. This is because the MSFC makes use of the scanline technique to divide the range blocks. Therefore, the high vertical-frequencies of the frame-based images are preserved. Furthermore, the MSFC greatly achieves the encoding and decoding time when compared with the CFC codec as demonstrated in Figure 3.10. We can summarize that the mutual domain pool design leads to a trade-off between the acceptable degradation of reconstructed image and the significant improvement of mutual fractal coding.

Although the mutual scanline fractal coding can greatly improve the encoding and decoding time, some limitations have still remained. Especially, the bit rate is still a little high. In addition, the structure of coding system is not generic. However, its concept is a basis approach for extending to the generic structure of coding system. This concept is introduced in Chapter 4.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Chapter 4

Mutual Fractal Coding: A Two-dimensional Approach

4.1 Introduction

As introduced in Chapter 3, the mutual scanline fractal coding for frame-based images is successful not only for reducing encoding and decoding time, but also for preserving high-vertical frequencies, however its coding efficiency is not satisfied for a very-low-bit-rate. In this chapter, the two-dimensional approach is proposed so that coding efficiency in terms of bit rate is improved.

The rest of this chapter is organized in the following way: In Section 4.2, the designing phase is introduced. The structure of MSFC is analyzed and is responsible for the unsatisfactory performance. Then the two-dimensional approach of mutual fractal coding is proposed. The implementation scheme is described in Section 4.3. The experimental results are illustrated in Section 4.4. Finally, the conclusion is given in Section 4.5.

4.2 Design of Coding System

All designing phases are almost the same as in the mutual scanline fractal coding case described in Chapter 3, therefore in this section the limitation of mutual scanline fractal coding approach is first analyzed, especially coding efficiency in terms of bit rate. The construction of mutual 2D-fractal coding (M2DFC) is proposed and the steps of design are described in the following subsections.

4.2.1 Analysis of Mutual Scanline Fractal Coding

This subsection begins with the simple analysis of mutual scanline fractal coding. Suppose the original image with size 256x256 and 8 bpp is considered, the smallest and largest range blocks fix the limits of the possible bit rate. As introduced in Chapter 3, the smallest and largest suitable range blocks are 1x8 and 1x32, respectively. For simple analysis the smallest range block is 1x8. In this case, its bit

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

rate that computed by means of Eqs.(3.18) and (3.21) is 2.875 bpp (23 bits/8 pixels). This is the upper boundary of the bit rate. For the lower boundary, the largest range block is 1x32, so its bit rate is 0.719 bpp. At this point, it can be summarized that the bit rate of the MSFC does not satisfy a very-low-bit-rate situation. This is because the largest range block is not large enough. On the other hand, we consider the two-dimensional (2-D) range block. The smallest range block is 4x4, hence its upper bound bit rate is 1.9375 bpp (31 bits/16 pixels). In this case, the bit allocation for each range block is based on the conventional fractal coding [39]. Similarly, if the largest range block is 16x16, then its lower bound bit rate is 0.1211 bpp (31 bits/256 pixels). It is found that the two-dimensional coding achieves the very-low-bit-rate situation. The two-dimensional range block is a generic structure for coding images and for this reason, the mutual fractal coding is extended to the 2-D case. The following subsections introduce its encoder and decoder.

4.2.2 Encoder

In the encoder section, the structure of range and domain blocks is formed as depicted in Figure 4.1. The frame image is decomposed into two field images, odd and even fields, and with this structure, the range blocks are partitioned into non-overlapping blocks. The consecutive non-overlapping blocks in row orientation are called *row-range blocks*, as illustrated in Figure 4.1(a). The mutual domain pools are set up in the same way. The odd-field image is extracted the overlapping domain blocks for the range blocks in even-field image and vice versa (See Figure 4.1(b)). Domain blocks partitioned in the same row-range blocks are selected for computing domain-range mapping. In the encoding process, the row-range blocks are encoded from top to bottom. Domain-range mapping, composed of two parts, geometry and intensity parts, is defined by

$$\begin{cases} f_{2D} = \left(f_{M_{2D}} \circ f_{G_{2D}}\right)(x_l, y_l) = f_{M_{2D}}\left(f_{G_{2D}}(x_l, y_l)\right) \\ g_{2D} = \left(g_{M_{2D}} \circ g_{G_{2D}}\right)(x_r, y_r) = g_{M_{2D}}\left(g_{G_{2D}}(x_r, y_r)\right) \end{cases} \tag{4.1}$$

where subscript 2D denotes 2-D affine transformations. The geometry part is replaced with Eq. (4.2) defined as follows:

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

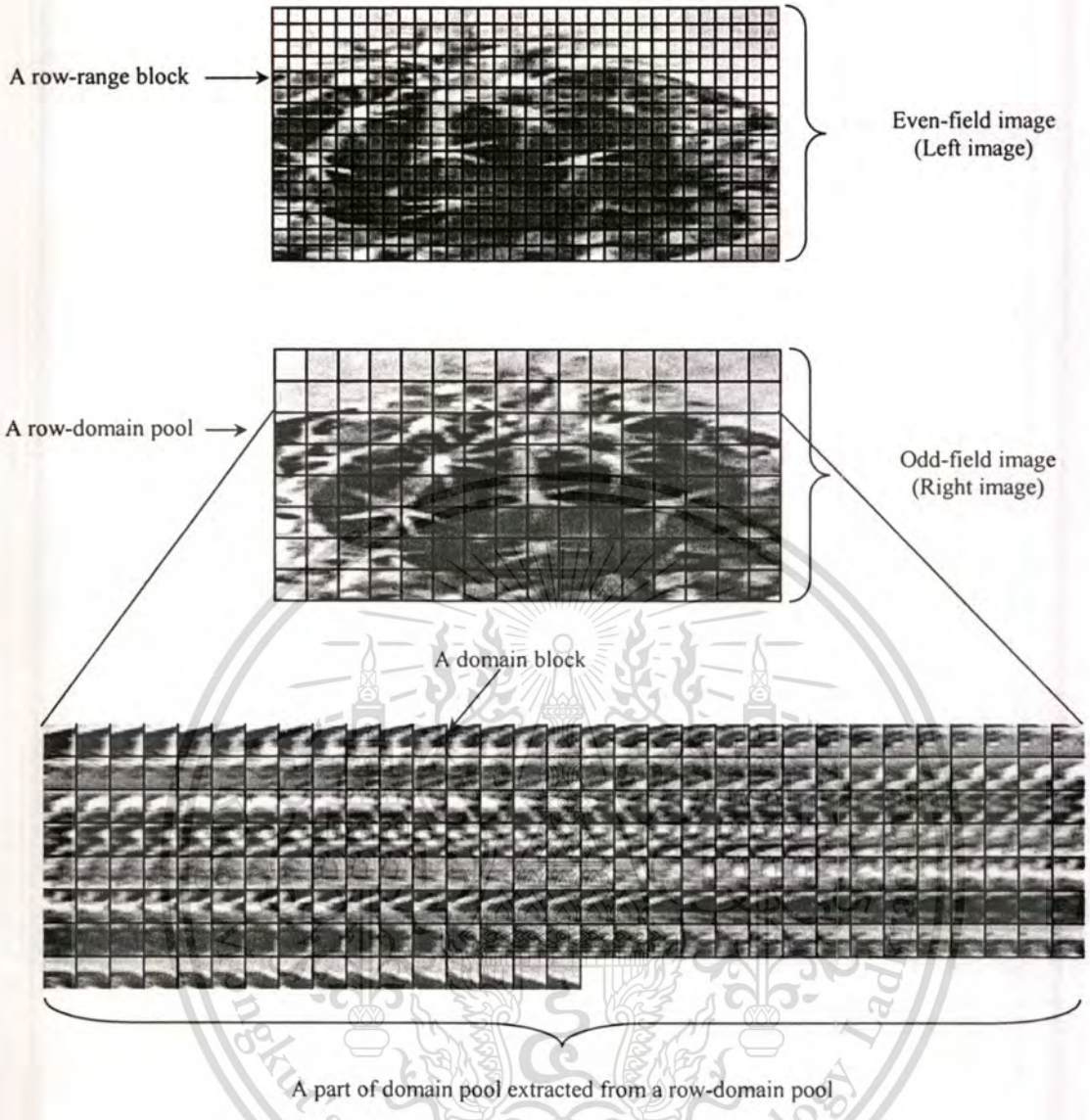


Figure 4.1 Structures of range blocks and domain blocks in a domain pool.

$$\begin{cases} f_{G_{2D}}(x_l, y_l) = \begin{bmatrix} l_1 & l_2 \\ l_3 & l_4 \end{bmatrix} \begin{bmatrix} x_l \\ y_l \end{bmatrix} + \begin{bmatrix} l_5 \\ l_6 \end{bmatrix}, & \forall x_l, y_l \in X_L \\ g_{G_{2D}}(x_r, y_r) = \begin{bmatrix} r_1 & r_2 \\ r_3 & r_4 \end{bmatrix} \begin{bmatrix} x_r \\ y_r \end{bmatrix} + \begin{bmatrix} r_5 \\ r_6 \end{bmatrix}, & \forall x_r, y_r \in X_R \end{cases} \tag{4.2}$$

where l_n and r_n denote the elements of a geometrical contractive matrix, and $[x_l\ y_l]^T$ and $[x_r\ y_r]^T$ represent the pixel vectors of the left and right images, respectively. For intensity part, the similarity between domain and range blocks is determined in the same way as the mutual scanline fractal coding approach.

This material is reserved for educational use only. Not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

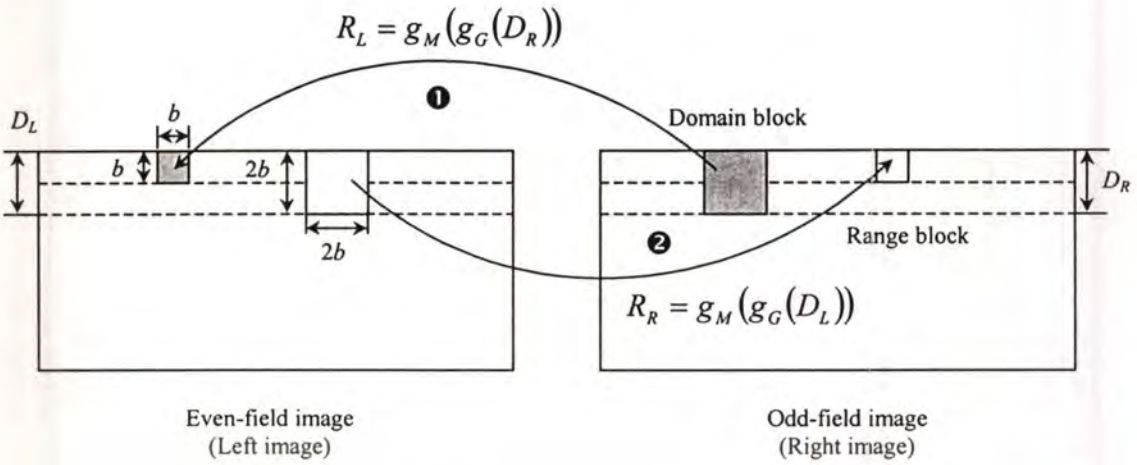


Figure 4.2 Decoding of mutual fractal codes.

4.2.3 Decoder

In order to describe the manipulation of the decoding process, the odd- and even-field images are set up with zero-initial values, i.e., all pixels of field images are zeroes. The seed images are the odd and even fields. For each iteration, there are two steps to decode the mutual fractal codes, as depicted in Figure 4.2. In the first step, all range blocks of the odd field are mapped with the appropriate domain-transformed blocks from the even field, such that $R_{L(v)}=g_M(g_G(D_{R(v)}))$, where $g_G(.)$ and $g_M(.)$ are contractive (geometric) and massic transformations and $D_{R(v)}$ is a domain block of the even filed. In the second step, the outcome in the odd field, which is a domain block of the even field, is performed with the same process as the first step. Note that the odd field, a seed image of the even filed, is not initialized with zero values, but holds the results from the first step. This implies that the seed image of the even field is initialized close to the original one, thus the transformation $g_M(f_G(.))$ converges rapidly to the fixed point. This is an effective method for generating the mutual fractal codes and can speed up the decoding time. In this way, the mutual fractal codes generated from the mutual domain-range mapping algorithms provide faster decoding of the reconstructed images.

4.2.4 Complexity Reduction of Mutual 2-D Domain Pools

According to the complexity reduction analyzed in subsection 3.4.6, it is clear that the complexity of the first-match-searching algorithm is $O(N_D)$. The domain

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

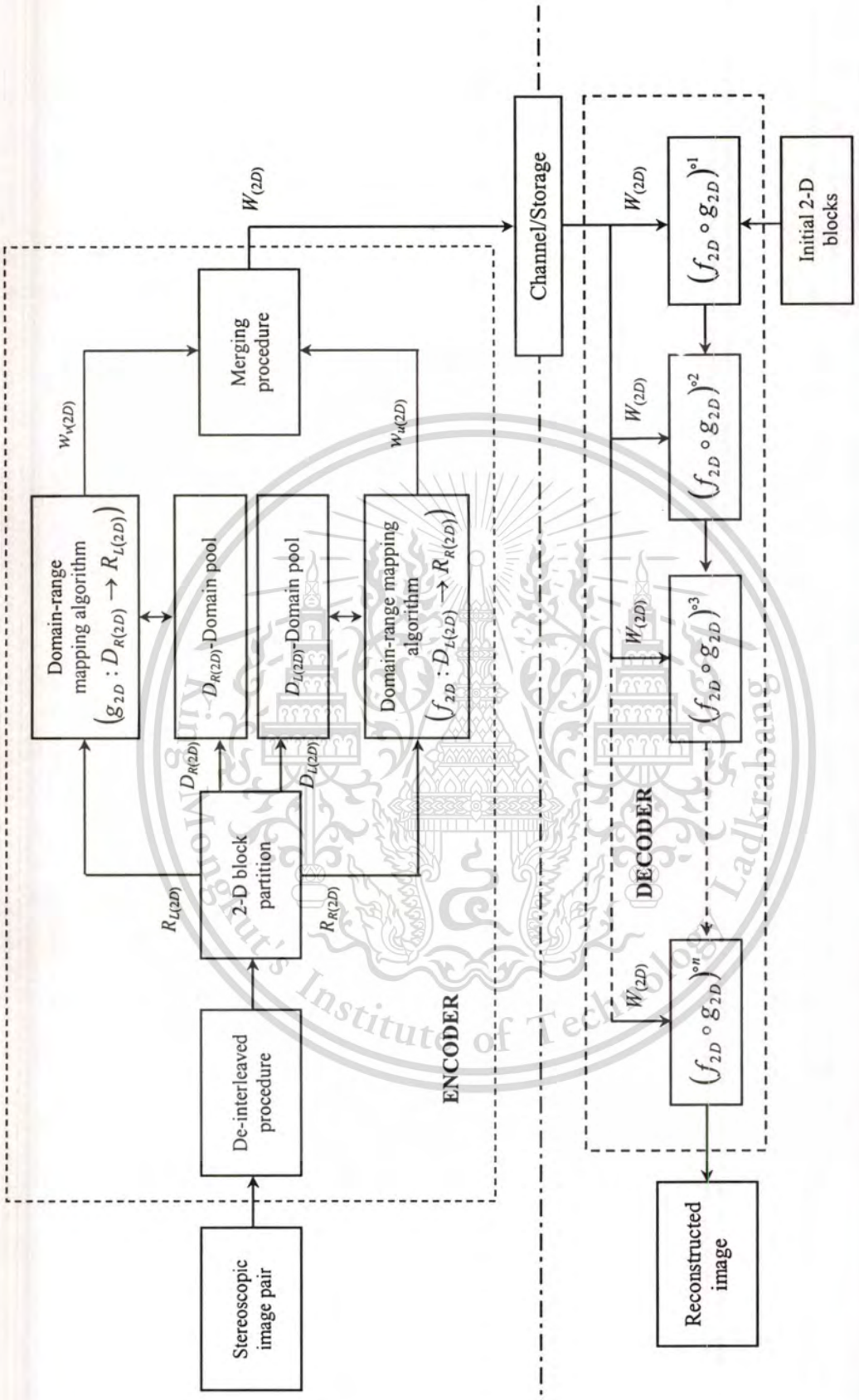


Figure 4.3 Overall implementation scheme of mutual 2-D fractal coding.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

pool size formed by the 2-D case is a principal factor, however, that makes the complexity reduction possible. Hence, the row-domain pool depicted in Figure 4.1 can be obtained by means of Eq. 3.20. It is found that the row-domain pool size is similar to the mutual scanline domain pool, i.e., $N_D = M\text{-}2R$, but the dimension of domain blocks is increased from 1-D to 2-D. This increases the computational time in domain-range mapping procedure, however when compared to the conventional one, the row-domain pools are still greatly reduced.

4.3 Implementation of Coding System

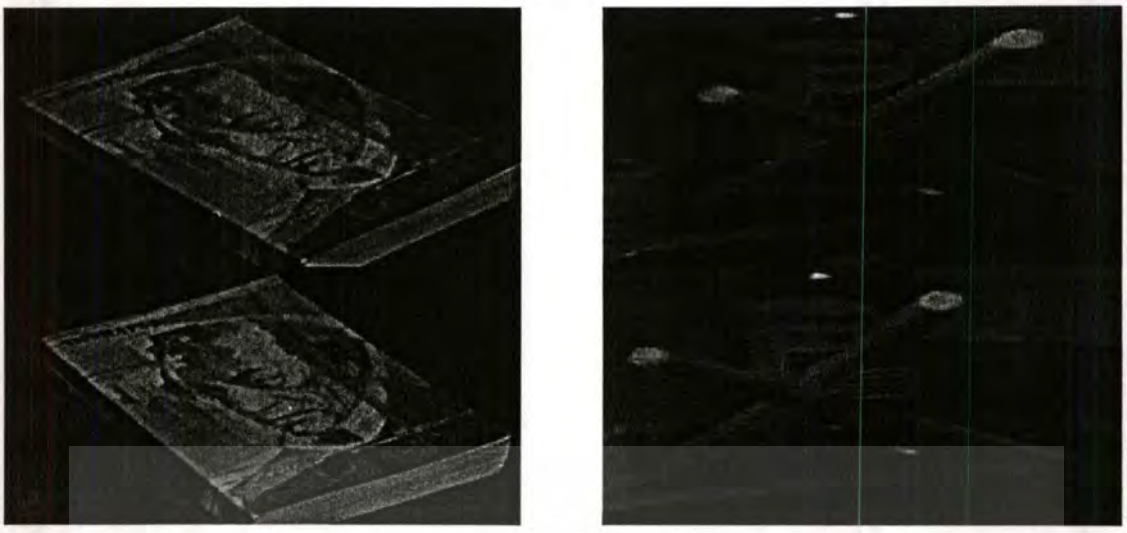
The implementation scheme of mutual 2-D fractal coding (M2DFC) is depicted in Figure 8(b). The main structure is similar to mutual scanline fractal coding (MSFC). All of designing phases described in Section 3.4 can be extended easily to the 2-D case, thus only the significant parts of the scheme are addressed in this chapter.

In this scheme, the preprocessing step of the encoder section is a de-interleaved procedure. This procedure converts a frame-based image, containing both even and odd fields, into a pair of field-based images shown in Figure 3.1(b) so that the high vertical-frequencies of the image are eliminated, then the field-based images are partitioned into 2-D domain and range blocks. The outcomes from the 2-D block partition are $R_{L(2D)}$, $R_{R(2D)}$, $D_{L(2D)}$, and $D_{R(2D)}$, where subscripts L and R represent the left and right images. The $D_{L(2D)}$- and $D_{R(2D)}$-domain pools are part of the field images and are independent, as shown in Figure 4.1. In this case, the operation of domain-range mapping algorithms performs the same as in the scanline case. The different parts are that Eqs. (3.14) and (3.21) are replaced with Eqs. (4.1) and (4.2), respectively. The bit requirement for mutual fractal codes of the M2DFC can also be obtained as follows:

$$B_{M2DFC} = D_{l(2D)} + D_{o(2D)} + R_{s(2D)} + s_{2D} + o_{2D} \tag{4.3}$$

where $D_{l(2D)}$, $D_{o(2D)}$, $R_{s(2D)}$, $s_{(2D)}$, and $o_{(2D)}$ denote domain-block location, domain-block orientation, range-block structure, scaling, and offset, respectively. In case of 256x256 image size, it totally needs 25 bits (2 bits for range-block structure (three-level quadtree partition: 4x4, 8x8, and 16x16), 8 bits for domain-block location

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



(a) (b)

Figure 4.4 Original stereoscopic test images in field-based format: (a) a "Book" image and (b) a "Chopstick" image.

(column coordinate only), 3 bits for domain-block orientation, 7 bits for offset, and 5 bits for scaling, respectively) for each self-transformation, i.e. mutual fractal code. It is clear in this way that the mutual-domain-pool design makes the coding system effective by the following reasons. (i) The domain-pool size is reduced, thus speeding up the search time while maintaining image quality as well as possible. (ii) Saving the number of bits for storing the domain-block position, only the column coordinate is kept, whereas the row coordinate is omitted, since it implicitly is with the range block position. (iii) The mutual fractal codes generated from the mutual domain-range mapping algorithms provide faster decoding of the reconstructed images for the reason in the Subsection 3.3.

4.4 Experimental Results

In order to evaluate the effectiveness of the mutual domain pool design in the 2-D case, the M2DFC codec is proposed and compared with the CFC codec. Prior to this experiment, the following conditions were set for the fairness of comparison. In case of the CFC codec, the specifications were not changed from the first experiment in Chapter 3. In case of the M2DFC, there were two versions of the range block partition, M2DFC-I for two-level quadtree partition (4x4 and 8x8 range-block sizes) and M2DFC-II for three-level quadtree partition (4x4, 8x8, and 16x16 range-block sizes), respectively. For M2DFC-I codec, the bit requirement is 24 bits (1 bit for

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

$R_{s(2D)}$), defined in Eq. (4.3), for each mutual fractal code. For M2DFC-II codec, the bit allocation is 25 bits (2 bits for $R_{s(2D)}$). The overall implementation scheme depicted in Figure 4.3 is applied to a set of test images, "Cactus", "Book" and "Chopstick", which are field-based images as shown in Figures 3.2(b) and 4.4, respectively. The experimental results show that the quality of the reconstructed images of the M2DFC-I, M2DFC-II and CFC codecs is comparable, as graphically exhibited in Figure 4.5. For high complex image, the average degradation of reconstructed image quality of the M2DFC-I is lower than 0.15 dB at 0.92 bpp when compared with the CFC codec, as illustrated in Figure 4.5(a). For the median complex image, it yields almost the same reconstructed images as the CFC codec. For the low complex image, it outperforms the CFC codec for a bit rate greater than 0.47 bpp. In the case of M2DFC-II, the average degradation is lower than 0.35 dB at 0.85 bpp for the high complex image and 0.15 dB at 0.49 bpp for the median complex image and 1.30 dB at 0.19 bpp for the low complex image, respectively, when compared with the CFC codec. Nevertheless, the encoding time of the M2DFC-I and M2DFC-II is greatly improved, as shown in Figure 4.6. The mutual fractal encoding time is faster than the conventional fractal coding, as summarized in Tables 4.1 and 4.2. In the case of decoding time, the experimental results illustrate that the M2DFC-I codec takes fewer iterations than the CFC codec when both codecs are encoded with the same scaling value (s=1). In Figure 4.7, the M2DFC-I codec outperforms the CFC in terms of taking fewer iterations; the M2DFC-I codec converges within about 3 to 4 iterations while the CFC needs more than 5 iterations to decode the reconstructed images. This implies that the intended codec is faster than the conventional one by approximately 2 times. This is a trade-off between the acceptable degradation of reconstructed images and the great improvement of mutual fractal encoding and decoding time.

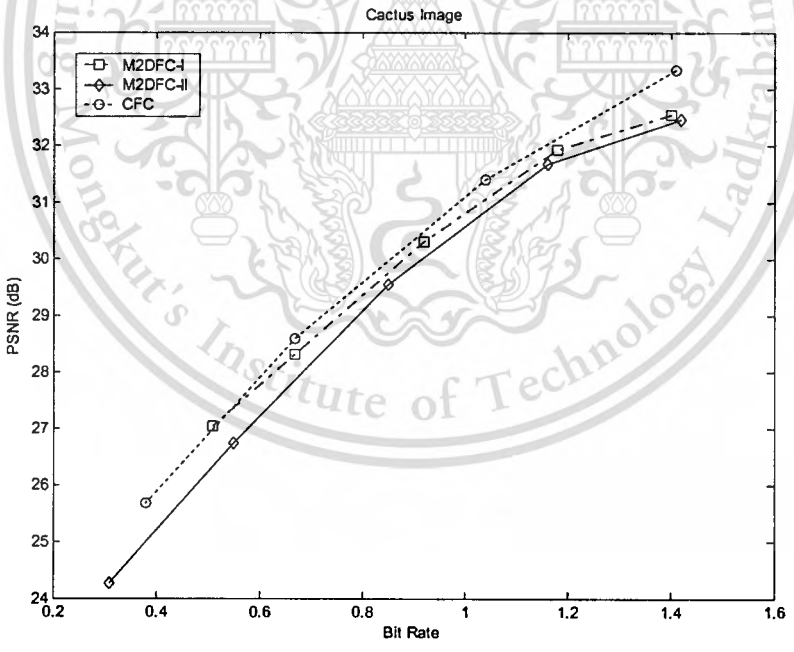
4.5 Conclusion

In this chapter, the extended concept of mutual domain pool design from a one-dimensional approach to a two-dimensional approach has been proposed. The objectives of this extended concept are to reduce the coding bit rate and to provide the coding structure to be a generic form. In order to achieve these objectives, the structure of domain and range blocks is formed under the following conditions. (i) The consecutive range blocks in the row orientation are grouped into a row-range

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

block. In the same way, the mutual domain pools are the collections of overlapping blocks. The consecutive domain blocks in the row direction are also formed into a row-domain block. (ii) Only the range and domain blocks, which are in the same row coordinate, are mapped and encoded with fractal codes in the encoding phase. These codes are also called *mutual fractal codes*. In this way, it is found that the coding bit rate is comparable with the conventional one. This is because the size of range blocks is larger while the D_y coefficient of affine transformation is omitted. Image quality is an acceptable degradation when compared with the CFC codec, but encoding and decoding time is greatly improved.

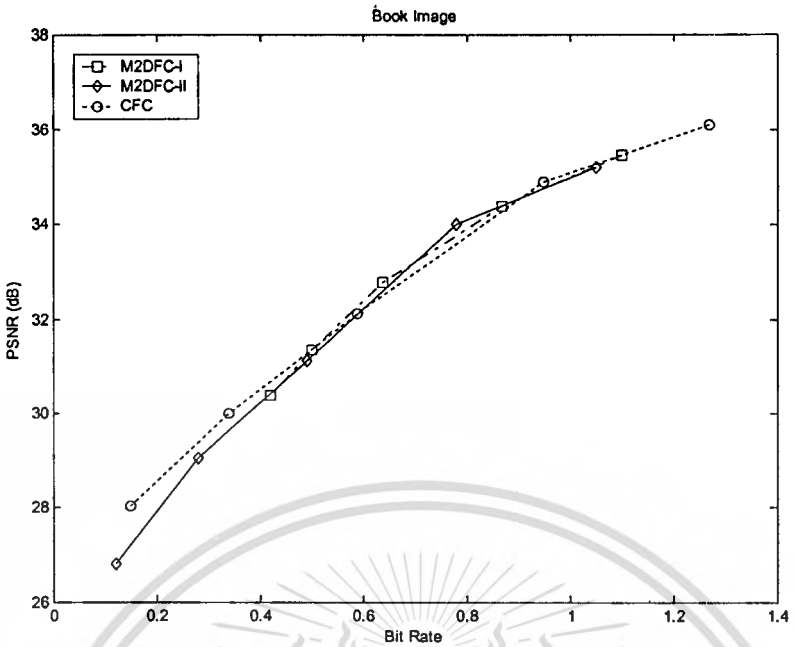
This satisfaction is not the end of developing mutual fractal coding. Chapter 5 introduces the combination method for improving its efficiency.



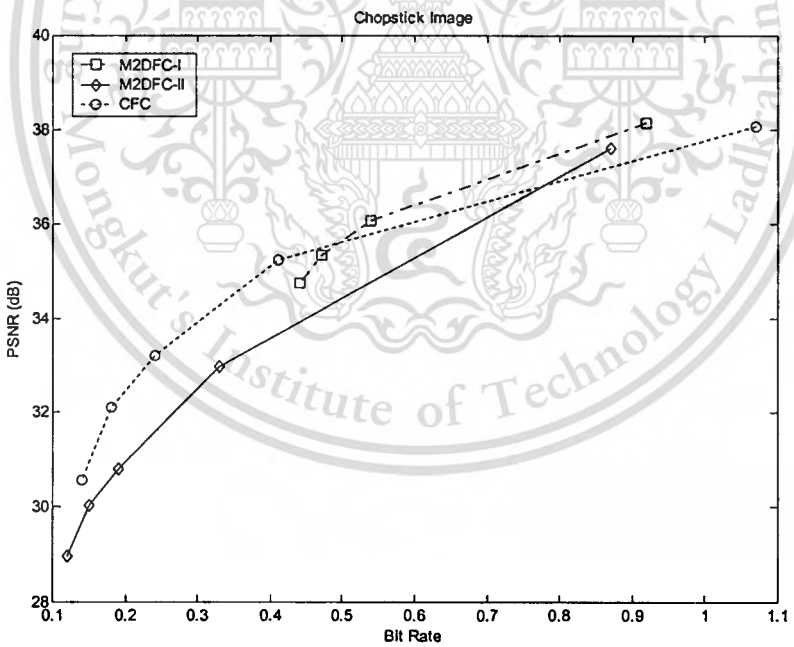
(a)

Figure 4.5 Comparison of the field-based reconstructed images of CFC with 10 iterations and M2DFC-I and M2DFC-II with 5 iterations.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



(b)



(c)

Figure 4.5 (Cont.)

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

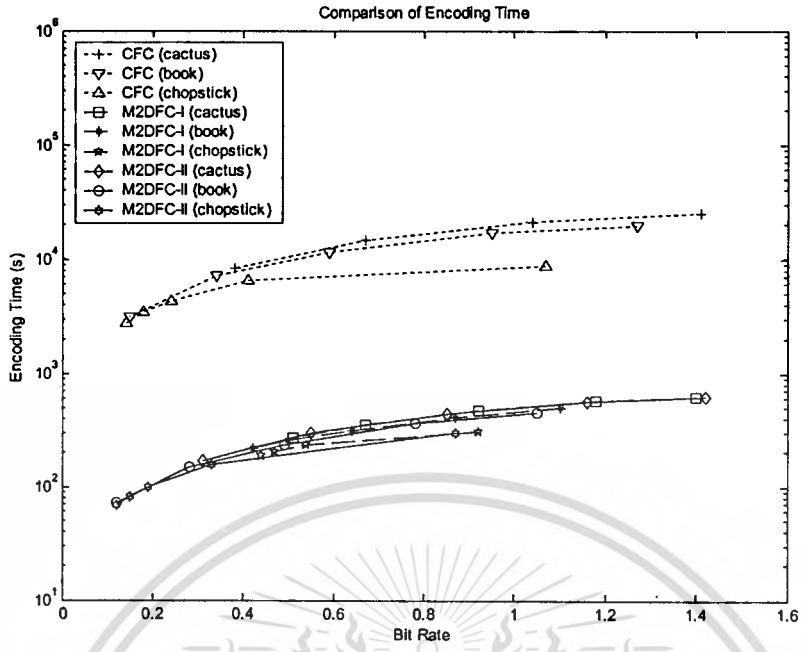
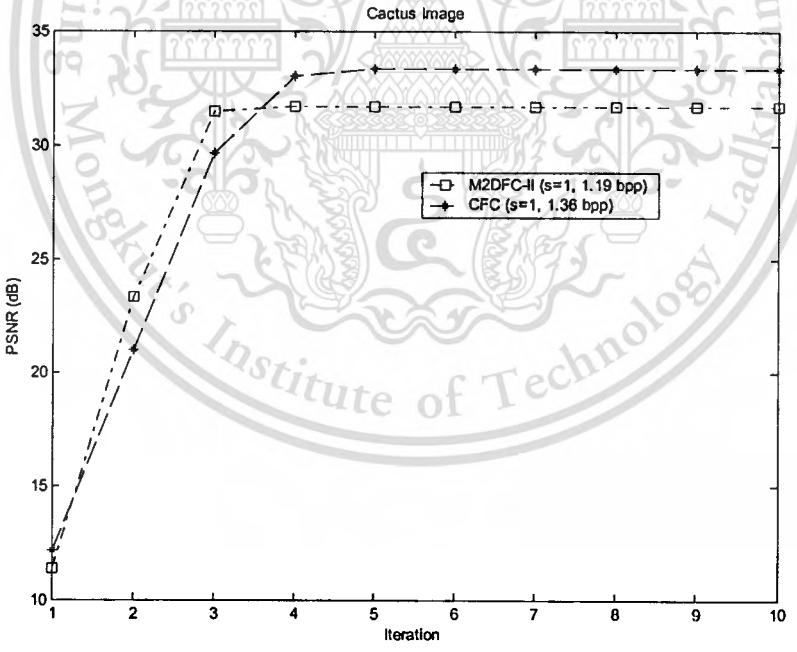


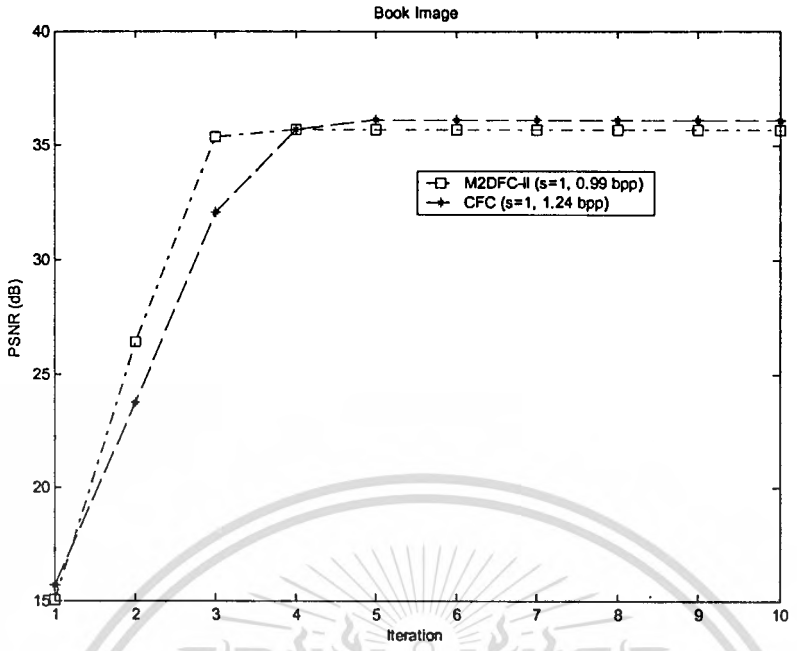
Figure 4.6 Comparison of encoding time of CFC, M2DFC-I and M2DFC-II in field-based coding.



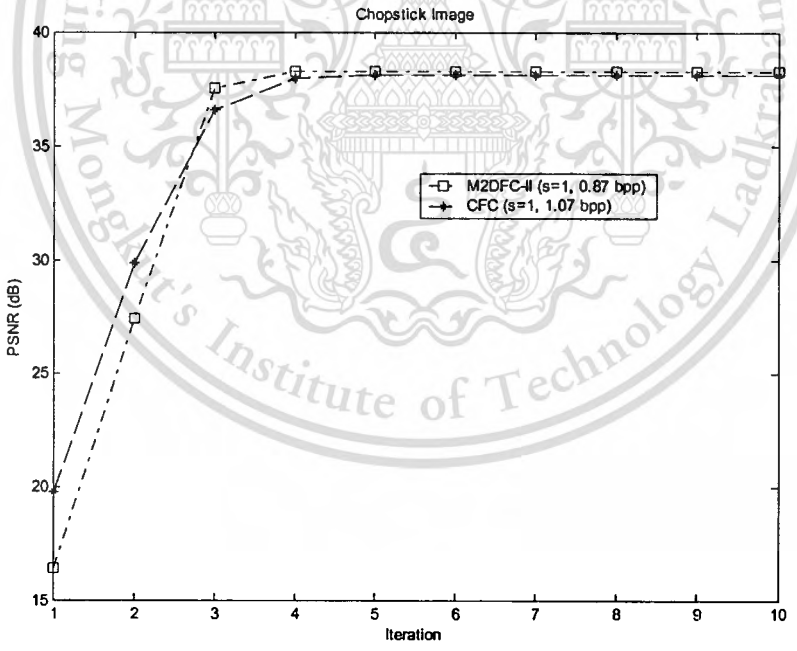
(a)

Figure 4.7 Comparison of the number of decoding iterations versus PSNR in field-basec coding.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



(b)



(c)

Figure 4.7 (Cont.)

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Table 4.1 Comparison of average encoding time of M2DFC-I and CFC.

Test image	M2DFC-I (s)	CFC (s)	Improved Ratio
Cactus	457.97	17,251.07	38
Book	337.54	11,676.23	35
Chopstick	231.01	5,137.30	22

Table 4.2 Comparison of average encoding time of M2DFC-II and CFC.

Test image	M2DFC-II (s)	CFC (s)	Improved Ratio
Cactus	417.86	17,251.07	41
Book	254.22	11,676.23	46
Chopstick	139.79	5,137.30	37

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Chapter 5

Mutual Fractal Coding Based on Variance

5.1 Introduction

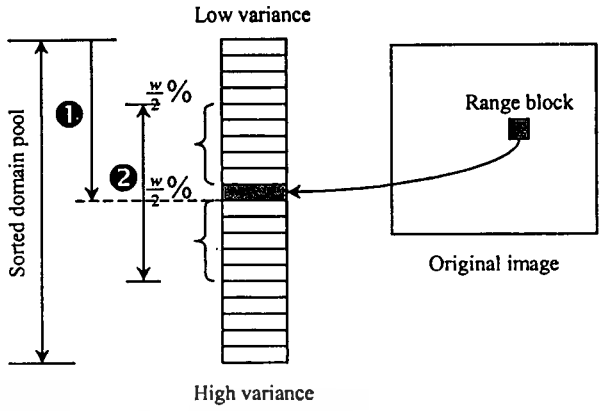
In this chapter, mutual fractal coding based on variance is proposed. This approach is a combination of the advantages of mutual fractal coding and variance-based fractal coding. The main objective of the intended method is to compromise the following factors: image quality, encoding time, decoding time and coding bit rate. The variance-based scheme proposed by [44], [45] is a simple and effective approach to improve fast encoding time and to achieve good image quality, but the decoding time and coding bit rate are still no worse than the conventional fractal coding. On the other hand, although the mutual fractal coding introduced in Chapters 3 and 4 can provide for both less decoding time and lower coding bit rate, the encoding time still has to speed up. In order to accomplish the objective, therefore, the mutual domain pool design is implemented for variance-based scheme.

The rest of this paper is organized as follows: Section 5.2 reviews the variance-based approach for speeding up the fractal image encoding. In Section 5.3, the mutual domain pool design is also addressed, then mutual fractal image coding based on variance is proposed in Section 5.4. The experimental results are illustrated in Section 5.5. Finally, the conclusion is given in Section 5.6.

5.2 Variance-based Fractal Coding

The variance-based approach is one of the most recent techniques used to speed up the fractal image encoding process. The main idea of this approach is based on the fact that two equal-sizes of image blocks cannot be closely matched unless their means and variances are closely matched [44]. Using this idea, it can concisely describe the practical procedure as follows. First, the original image is partitioned into non-overlapping range blocks, then domain blocks are extracted from the original image with twice size of range blocks. A collection of domain blocks is called a *domain pool*. Next, each domain block in the domain pool is contracted to the range block size by pixel averaging and its variance is calculated. All the variances of the

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



❶ Searching the closest variance matching between domain and range blocks.
❷ Assigning the symmetrical searching window size.

Figure 5.1 Symmetrical searching window size and its operation.

domain blocks are sorted. The domain pool whose variances are sorted is called a *sorted domain pool*. After providing the range blocks and sorted domain pools, it reaches to the phase of obtaining fractal code or fractal image encoding. In this phase, each range block searches for the candidate of a matching domain block in the sorted domain pool. If the variances of range and domain blocks are closely matched, then the location of that domain block is marked to be a center or reference point of the searching window. The closest matching variance between range and domain blocks usually does not imply that they are matched in intensity affine transformation (massic transformation). For this reason the sorted neighboring variances are also included to calculate the optimal coefficients of intensity affine transformation by means of symmetrical searching windows, as illustrated in Figure 5.1. Finally, all the coefficients of transformation are stored in the form of fractal codes.

Recently, an improved variance-based scheme was proposed by C. He *et al* [45]. The main objectives of this scheme are to accelerate the encoding time and to improve the reconstructed image quality to be better than the original one. C. He introduced the improved scheme by appending a simple classification and a reduced domain pool criterion to the preprocessing procedure. In the classification phase, the range blocks are grouped into shaded and non-shaded blocks by the use of variance defined as

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

$$Var(R) < \varepsilon \tag{5.1}$$

where ε and R denote variance threshold and range block, respectively. If the variance of each range block is lower than a predefined threshold ε, then it is regarded as a shaded block, otherwise it is a non-shade block. Then all shaded blocks are encoded with their intensity means, whereas all non-shaded blocks are encoded with an intensity affine transformation. In addition, the domain pool reduces its size by means of the following criterion defined as

$$\Omega_\delta = \{D \in \Omega \mid Var(D) \geq \delta^2\} \tag{5.2}$$

where $\Omega, \Omega_\delta, D,$ and δ represent a domain pool, reduced-domain pool, domain block and reduced-domain-pool threshold, respectively. In this case, only the relatively large variances in the sorted domain pool are considered in order to obtain a smaller contrast factor s in intensity affine transformation. This makes the encoding time shorter and the reconstructed image improved, when compared with the variance-based scheme proposed in [44].

5.3 Mutual Domain Pool Design

Domain pool design for fractal image compression is an important phase to reduce the computational encoding time, therefore in this section, the mutual domain pool design is proposed to achieve not only less encoding time, but also fewer iterations of decoding time and coefficients of geometric affine transformation. This is a principal concept, which differs from the variance-based and conventional approaches. The designing phases are described in the following subsections.

5.3.1 Analysis of Conventional Fractal Coding

This subsection begins with the simple analysis of conventional fractal image compression. The affine transformation, which is composed of geometric and intensity parts, is widely used for fractal image compression. In fractal encoding process, the original image has to be partitioned into two types of subimage blocks, namely, range and domain blocks, then the affine transformation is applied to obtain the optimal coefficients of domain-range mapping. These coefficients are called *fractal codes*. The fractal codes are also composed of geometric and intensity coefficients. The original image with size 256x256 and 8 is considered here. The

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

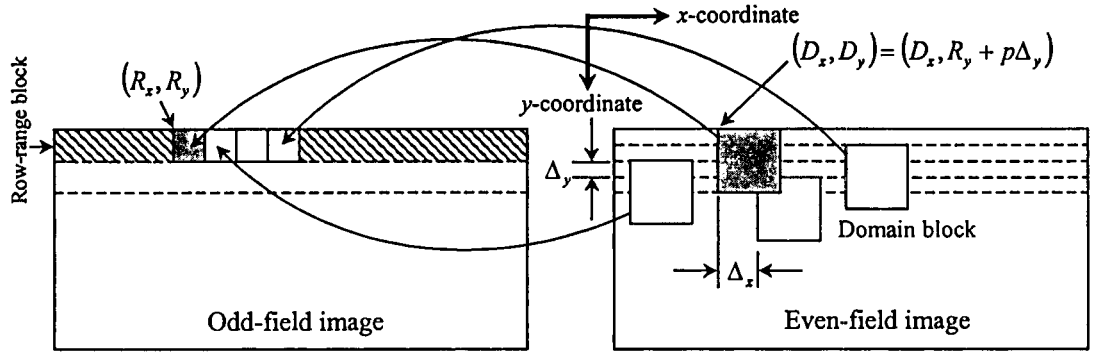


Figure 5.2 Mutual domain pool structure with minimizing the *y*-coordinate of domain blocks.

equal-size image partition is used for simple analysis. Bit requirements for this image can be classified as follows. In the geometric part, the coefficients are composed of *x*- and *y*-coordinates, D_x and D_y, of domain blocks, including isometric parameter (rotation and flip) I. In the intensity part, the coefficients consist of offset o (brightness) and scaling s (contrast). The bit allocations of these coefficients are 19 bits for geometric coefficients (16 (8+8) bits for domain block position and 3 bits for rotation and flip operations) and 12 bits for intensity coefficients (7 bits for brightness and 5 bits for contrast). Each range block needs 31 bits, i.e., the fractal code is tabulated as $(D_x, D_y, I, o, s) = (8, 8, 3, 7, 5)$. At this point, it is can be summarized that only the coefficients of domain block position take more than 50% of the amount of bit allocations for encoding each range block. In addition, the larger images need more bit requirements for encoding the coordinate coefficients. It is evident therefore, that the appropriate structure of the domain pool design can provide less encoding time and lower coding bit rate.

5.3.2 Modified Structure of Mutual Domain Pools

The main objectives of this designing phase are to reduce the bit requirement of coefficients of geometric affine transformation and the domain pool size. In order to accomplish these objectives, the mutual domain pool design is proposed. In this method, the original image is firstly de-interleaved into odd and even fields, so that the similarity characteristic of images still remains. Now the even and odd fields can be viewed as two equivalent complete metric spaces that the contractive mapping fixed point still holds [37]. This means that the domain pool of the original image is

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

also decomposed into two parts. The geometric of domain pools can be formed as illustrated in Figure 5.2. In this case, the structure of the range blocks is considered. The consecutive range blocks in row direction of the image are formed and called *row-range blocks*. Each range block of the row-range block is encoded in order from left to right and from top to bottom, respectively. Therefore, R_x and R_y, which denote the range-block position, can be omitted, but the image size and range-block size are sent to the decoder by file's header for recovering the range block position. In a like manner, a collection of consecutive domain blocks extracted from the mutual domain pools is called a *row-domain pool.* This structure differs from the conventional fractal coding in that the domain pool for row-range blocks in the odd-field image is extracted from the even-field image and vice versa. This directly influences the number of iterations of decoding process explained clearly in the Subsection 4.2.3. In order to reduce the bit requirement of the coefficients of geometric affine transformation, we replace the y-coordinate of the domain pool with two parameters, Δ_y and p. Δ_y is defined as a step size increment of the row-domain-pool location in the y-coordinate direction. Although this parameter is not a part of fractal codes, it is important for decoding the domain block position, so it is attached to the file's header. p is a binary code to represent the number of row-domain pools. As exhibited in Figure 5.2, the domain pool size can be obtained by

$$D_s = \frac{2^p M - 2^{p+1} R}{\Delta_x} \tag{5.3}$$

where p, M, R, and Δ_x represent a number of row-domain pools, column image dimension, range block size, and horizontal step size increment, respectively. By this structure, only the x-coordinate of the domain block is kept, whereas the y-coordinate is replaced with $R_y+p\Delta_y$, since it is implicitly within the range block position. This makes the number of bits for storing the domain-block position saving. The domain pool size is also reduced, thus speeding up the encoding time.

In this way, it is clear that the mutual-domain-pool design makes the coding system effective by the following reasons. First, the domain-pool size is reduced by parameter p, thus speeding up the searching time, while maintaining the image quality as well as possible. Second, the number of bits for coefficients of the mutual fractal codes is reduced by D_y coefficient. This leads to a lower coding bit rate.

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

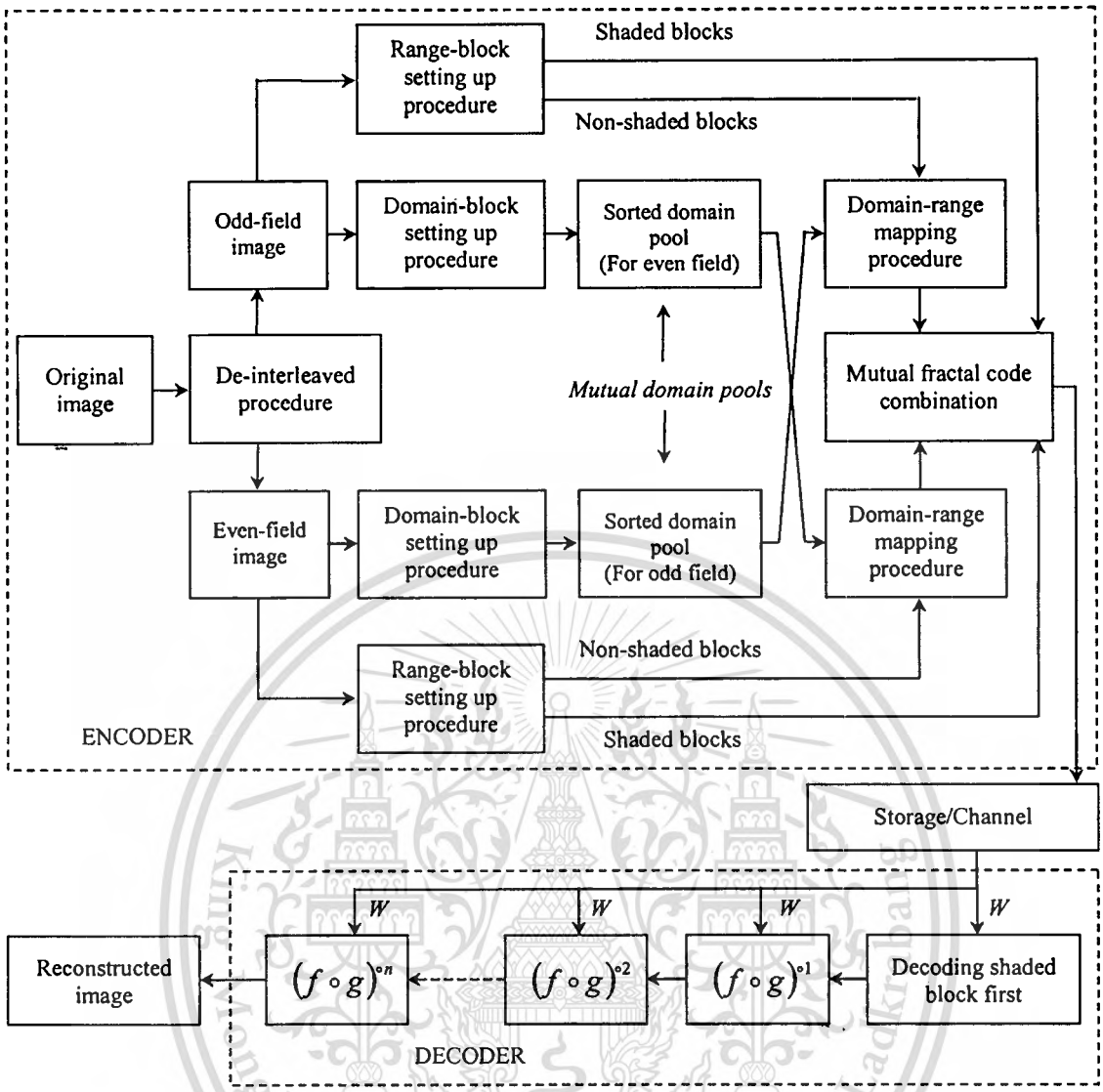


Figure 5.3 Overall implementation scheme of the proposed method.

5.3.3 Complexity Reduction of Modified Mutual Domain Pools

As described in Subsection 5.3.2, the objective of the modified structure is to extend the feasibility of self-similarity area to the vertical direction. This modified structure is able to improve the quality of reconstructed images, however the extension of self-similarity area leads to the enlargement of the mutual domain pool size. In order to limit the self-similarity area for an appropriate situation, i.e., compromise between complexity and reconstructed image quality, the following parameters, p and Δ_x, in Eq. 5.3 have to be controlled. The extension of the self-similarity area is p, whereas the limitation of mutual domain pool size is Δ_x. In this case, the complexity

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

of this structure is analyzed by setting p and Δ_x to 3-bit and 2-step increment, respectively. The mutual row-domain pools are entirely equal to 4(M-2R), which is four times greater than the original mutual-domain-pool size.

5.4 Proposed Approach

The overall implementation is schematically illustrated in Figure 5.3. The operation of this scheme can be described as follows: beginning with the encoder section, first, the original image is de-interleaved into odd- and even-field images, then each field image is partitioned into non-overlapping range blocks by a range-block setting up procedure. This procedure provides consecutive range blocks in the same row of the original image grouped into a row-range block. In addition, the elements of each row-range block are classified into two types of blocks, shaded and non-shaded blocks, as introduced in Section 5.2. All of shaded blocks are encoded by their intensity-mean values, while all of non-shaded blocks are encoded with coefficients of intensity affine transformation. The mutual domain pools are setup by a domain-block setting up procedure. In this procedure, the domain blocks are extracted from the odd- and even-field images, respectively. The number of domain blocks is limited by Eq.(5.3). The important factor in this case is the parameter 2^p, which represents a number of row-domain-block levels as depicted in Figure 5.2. This implies that the more bits for parameter p, the more levels of mutual row-domain pool are included, then all domain blocks obtain their variances. Only the relatively high variances are possessed and sorted to provide for domain-range mapping procedure. At this point, it can be imagined that the domain pool size is reduced by two factors. The first is a mutual-domain-pool structure minimizing with parameter 2^p and the second is a variance parameter; only high variance domain blocks are considered, so the low variance domain blocks are ignored. After providing the row-range blocks of each field images and sorted mutual domain pools, the mutual fractal codes are generated by domain-range mapping procedures. Finally, The outcomes from these procedures are combined together.

For the decoder section, the shaded blocks encoded by intensity-mean values are firstly decoded, then the mutual fractal codes are decoded by an iterative system function $(f \circ g)$, as illustrated in Figure 5.3. Finally, after a few iterations, the reconstructed image appears.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

5.5 Experimental Results

In this experiment, the performance of the intended codec is tested and evaluated in following issues: (i) image quality, (ii) encoding time, (iii) decoding time and (iv) coding bit rate. The condition of comparisons of the codec performance is based on the fractal coding on a spatial domain. The mutual-2D-fractal coding (M2DFC), variance-based fractal coding (VBFC) [45], conventional fractal coding (CFC) [39], and proposed method (MFCBV-I, -II, and -III) algorithms are implemented and simulated by using MATLAB programming. The machine for performing the simulation program is the desktop PC with PENTIUM 4 processor and the clock speed is 2.4 GHz.

In order to test and evaluate the performance of the intended codec compared with conventional codec, mutual-2D-fractal codec and the state-of-the-art of variance-based codec, a set of stereoscopic test images have been prepared in field-based format, i.e., "Cactus", "Book" and "Chopstick" images, as shown in Figures 3.1(b), 4.4(a) and 4.4(b), respectively. The image quality is evaluated by means of peak signal-to-noise ratio (PSNR), as defined in Eq.(3.20). For coding efficiency, evaluation is in the form of bit rate B_R, that can be calculated by Eq.(3.21).

For the first experiment, the following conditions are set for the fairness of comparison. In the case of the CFC and M2DFC-I codecs, the specifications are not changed from the first experiment in Chapter 3 and Chapter 4, respectively. In case of the VBFC and MFCBV-I codecs, the bit requirement for each fractal code is allocated in Table 5.1. These codecs also need a special parameter for identifying the range block classification, i.e., shaded- and non-shaded blocks. That parameter is C_g requiring only 1 bit. The additional codec parameters are also predefined as shown in Table 5.2. The parameter p in Eq.(5.3) is set to zero so as to minimize the mutual domain pool as small as possible. Searching window sizes of the VBFC and MFCBV-I codecs are set to 5% of entire domain pool and 50% of row-domain pool, respectively, then the overall implementation scheme depicted in Figure 5.3 is applied to a set of test images. The experimental results show that the quality of reconstructed images of the CFC, M2DFC-I, VBFC and MFCBV-I codecs is comparable as graphically exhibited in Figure 5.4. For the high complex image, the proposed and CFC codecs are almost the same quality of reconstructed image and outperforms both the M2DFC-1 and VBFC codecs. For the median complex image, the reconstructed

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

image quality of the proposed codec is superior to that of both M2DFC-I and VBFC codecs, but there is a little degradation when compared with the CFC codec. For the low complex image, MFCBV-I codec outperforms the VBFC codec for a bit rate greater than 0.85 bpp, whereas the average degradation is lower than 0.30 dB at 0.50 bpp when compared with the CFC codec. At this point, it can be summarized that the proposed method responds well in terms of efficiency of reconstructed image quality.

In addition, the experimental results illustrate that the MFCBV codec greatly improves the encoding time when compared with the CFC codec and compromises between the M2DFC-I and VBFC codecs, as graphically shown in Figure 5.5. (See also Tables 5.1 and 5.2)

In case of decoding time, the experimental results illustrate that the mutual fractal coding approach (M2DFC-I and MFCBV-I codecs) takes fewer iterations than the CFC and VBFC codecs when all of the codecs is encoded with the same scaling value (s=1). In Figure 5.6, the M2DFC-I codec outperforms the CFC codec in terms of taking fewer iterations; the MFCBV-I codec converges within about 3 to 4 iterations, while the VBFC codec needs at least more than 5 to 6 iterations to decode the reconstructed images. This implies that the proposed codec is faster than the CFC and VBFC codecs by approximately 2 times.

For the second experiment, the modified structure of the mutual domain pool designed in Figure 5.2 is implemented for the MFCBV-II and -III. The codec parameter values are set as Table 5.3. The searching window size of MFCBV-II and -III codecs is assigned as 5% and 10%, respectively. For the "Cactus" image, the experimental results show that the MFCBV-III codec is almost identical to the CFC codec and is superior to the other codecs in terms of reconstructed image quality (see Figure 5.6(a)). For the "Book" image, the quality of the reconstructed image is slightly degradable when compared with the CFC, M2DFC-I and MFCBV-I, but is better than the VBFC codec. For the "Chopstick" image, the MFCBV-III codec outperforms all of the codecs. On the other hand, the MFCBV-II codec gets the lowest ranking order in terms of reconstructed image quality for all test images. In case of encoding time, both the MFCBV-II and -III codecs are comparable to the VBFC codec and are superior to the others for all test images.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

At this point, it can be summarized that the modified structure of the mutual domain pool can provide good efficiency for both reconstructed image quality and encoding time, when suitable values are set for codec parameters, i.e., p, Δ_x, and Δ_y.

Table 5.1 Specifications of experimental codecs.

Codec	Bit allocation for encoding parameters							Total bit allocation
	D_x	D_y	I	o	s	R_s	C_g	
CFC (3-level)	8	7	3	7	5	2	-	32
M2DFC-I	8	-	3	7	5	1	-	24
VBFC	8	7	3	7	5	1	1	32
MFCBV-I*	8	-	3	7	5	1	1	25

* Note: Domain pool design is based on the approach described in Chapter 4.

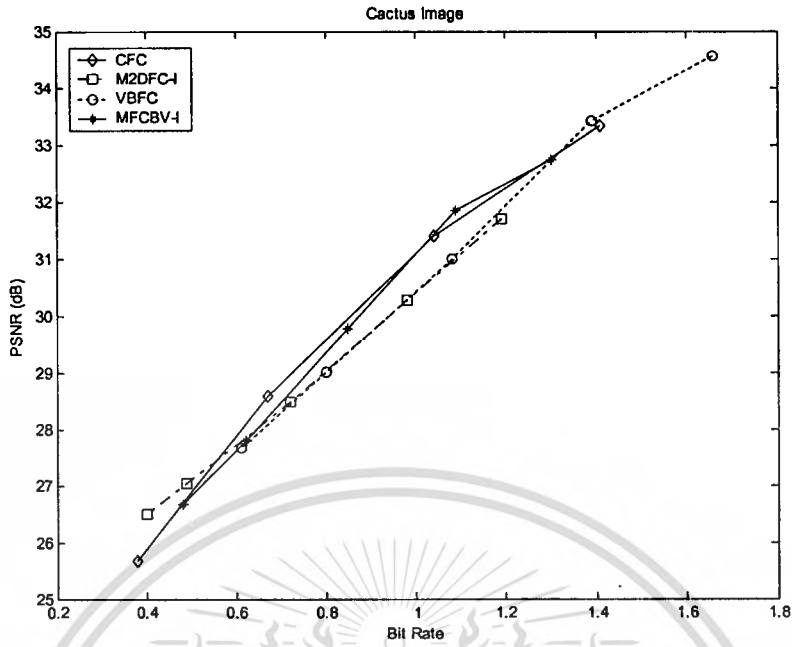
Table 5.2 Setting up predefined thresholds for VBFC, MFCBV-I, -II and -III codec parameters.

Codec parameters	Predefined thresholds
Variance threshold for classification (ε)	< 10 ; for non-shaded block
Variance threshold for domain pool reduction (δ)	≥ 10 ; for domain pool reduction
Variance threshold for matching block (η)	≤ 10 ; for domain-range matching

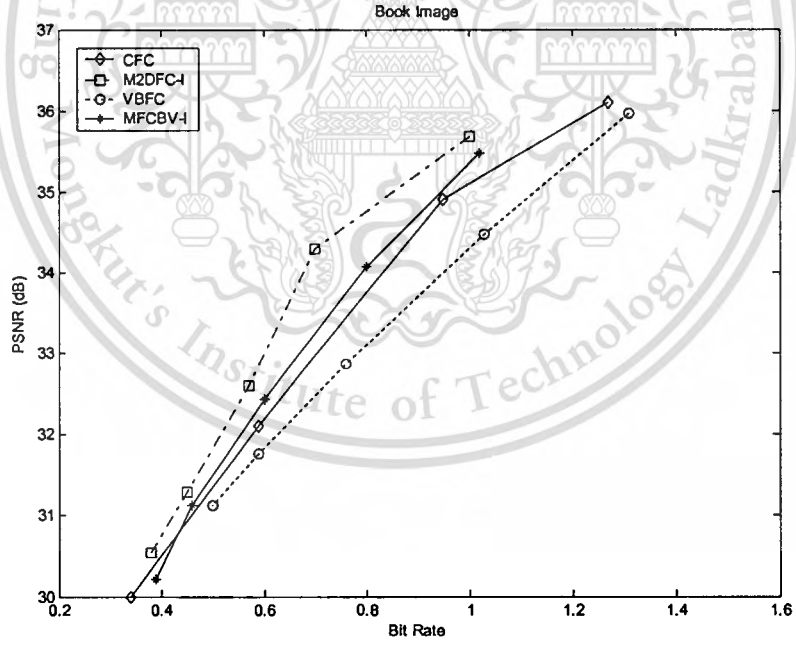
Table 5.3 Setting up parameter values for MFCBV-I and -II codecs.

Codec parameters	Parameter values
p-parameter	8-level (3 bits)
Δ_y	4-step increment
Δ_x	2-step increment

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



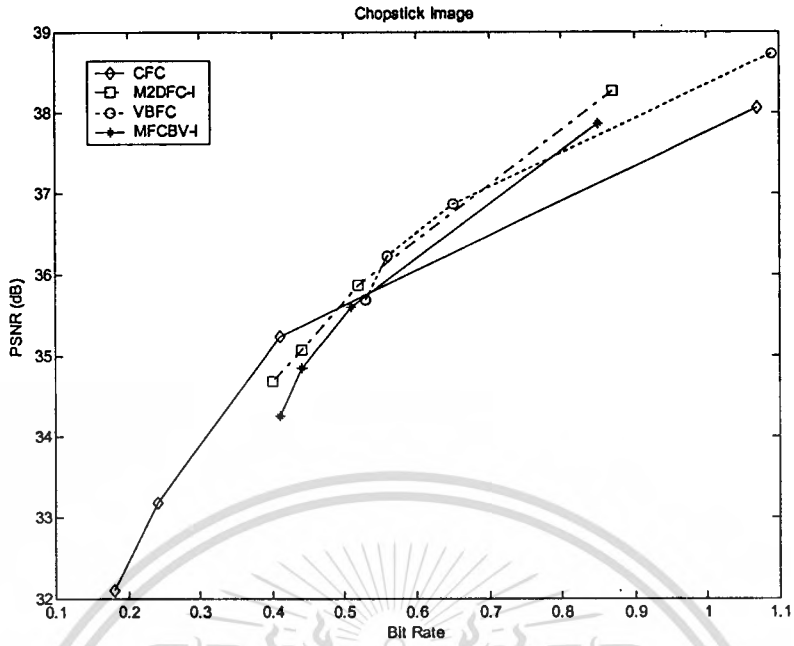
(a)



(b)

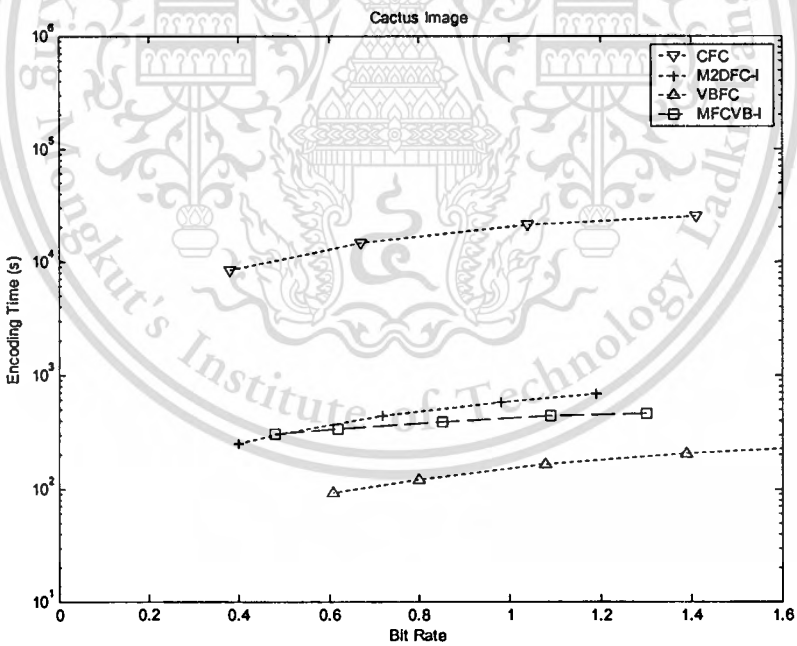
Figure 5.4 Comparison of the reconstructed field-based images of CFC, M2DFC-I, VBFC and MFCBV-I codecs.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



(c)

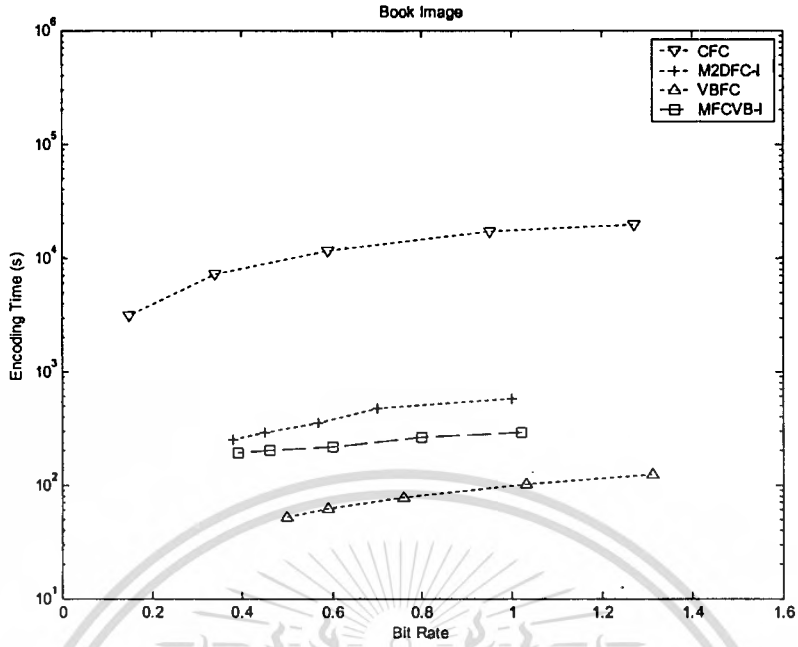
Figure 5.4 (Cont.)



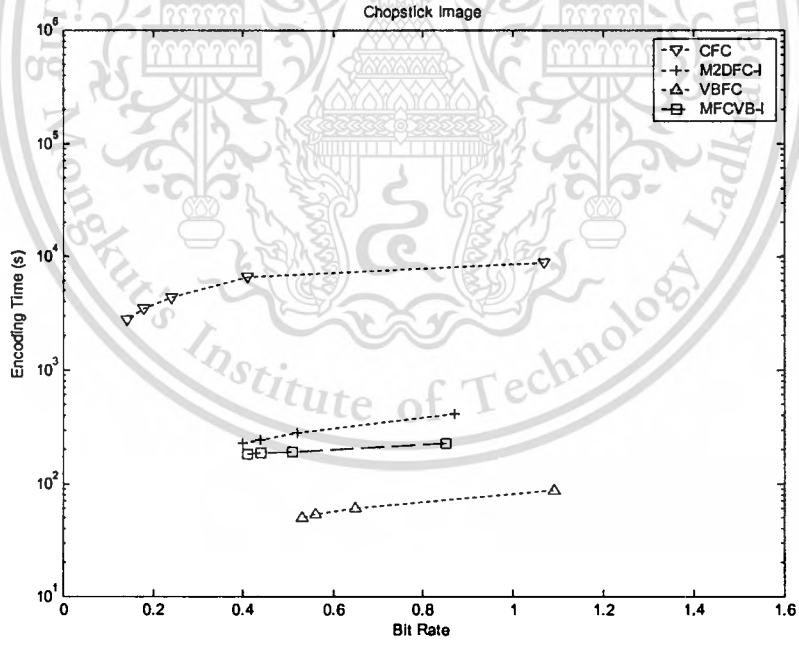
(a)

Figure 5.5 Comparison of encoding time of CFC, M2DFC-I, VBFC and MFCBV-I codecs.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



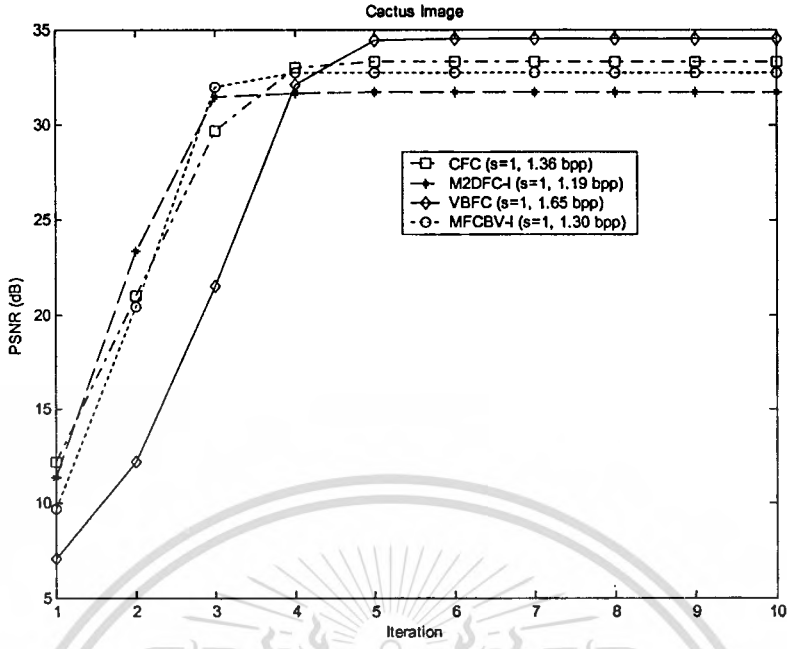
(b)



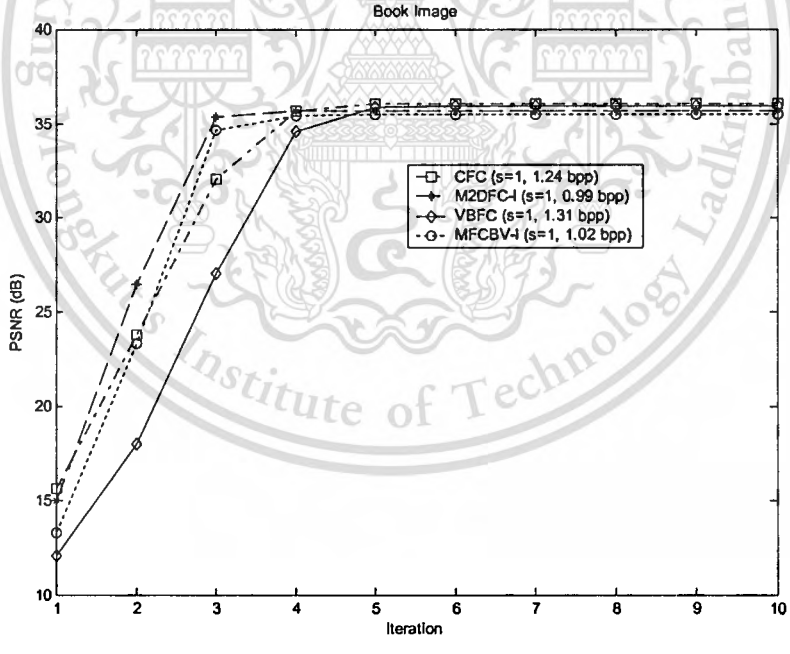
(c)

Figure 5.5 (Cont.)

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



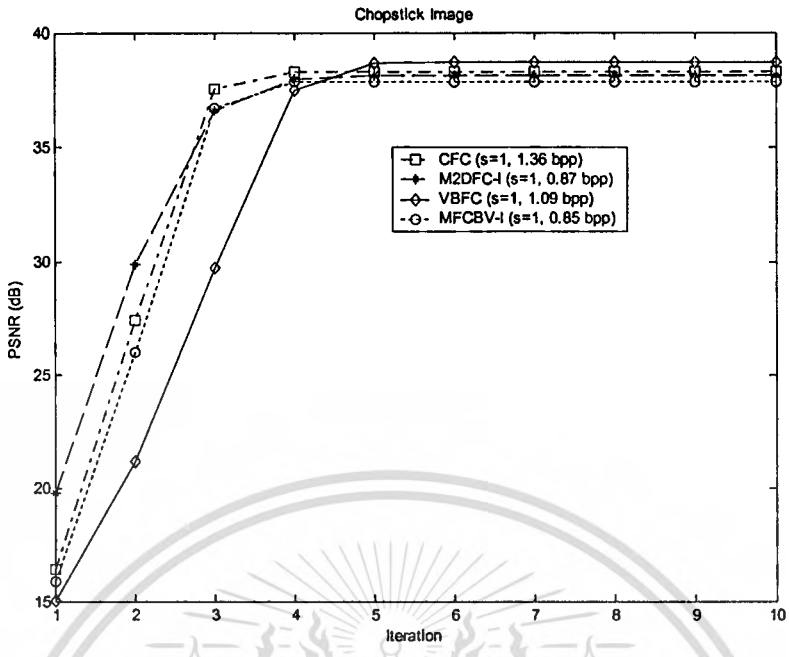
(a)



(b)

Figure 5.6 Comparison of decoding time of CFC, M2DFC-I, VBFC and MFCBV-I codecs.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



(c)

Figure 5.6 (Cont.)

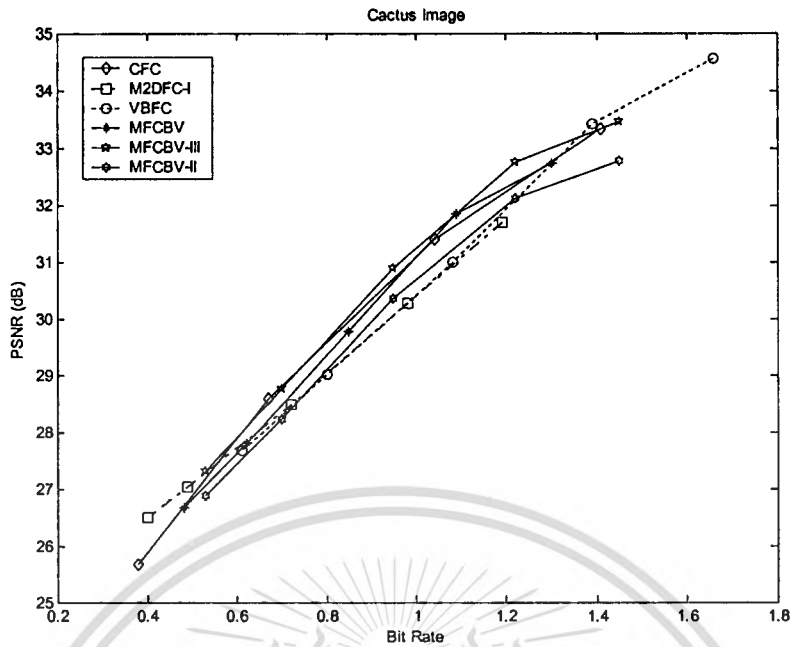
Table 5.4 Comparison of average encoding time of VBFC and CFC.

Test image	VBFC (s)	CFC (s)	Improved Ratio
Cactus	163.22	17,251.07	105
Book	83.00	11,676.23	140
Chopstick	60.03	5,137.30	85

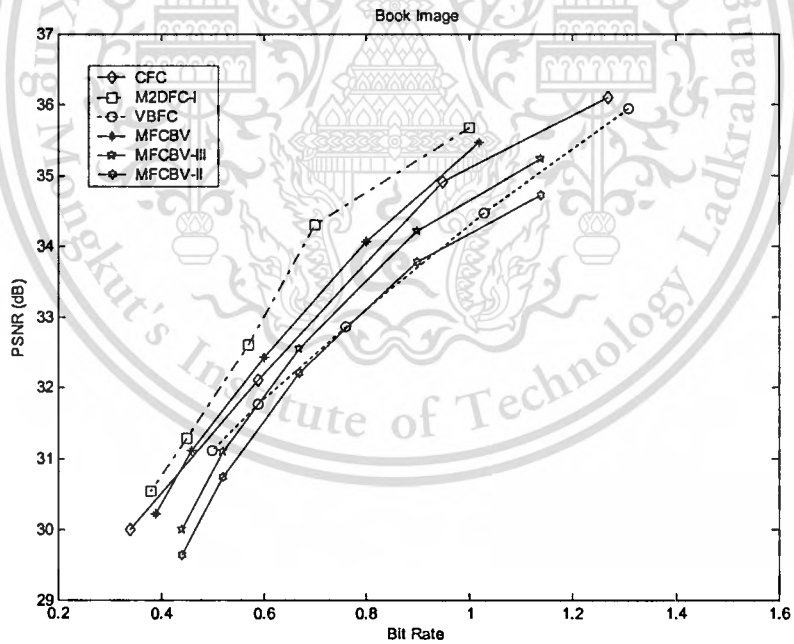
Table 5.5 Comparison of average encoding time of MFCBV and CFC.

Test image	MFCBV (s)	CFC (s)	Improved Ratio
Cactus	386.86	17,251.07	44
Book	232.32	11,676.23	50
Chopstick	195.53	5,137.30	26

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



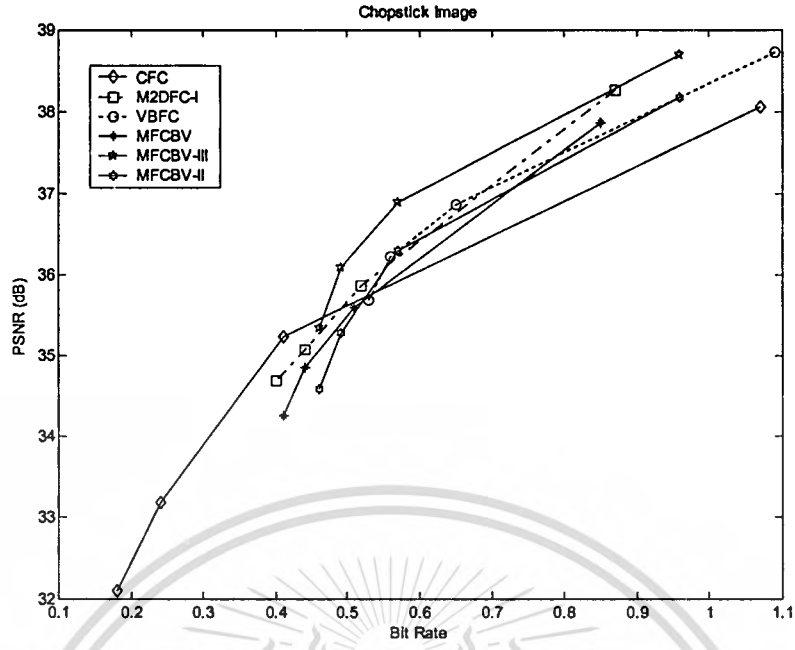
(a)



(b)

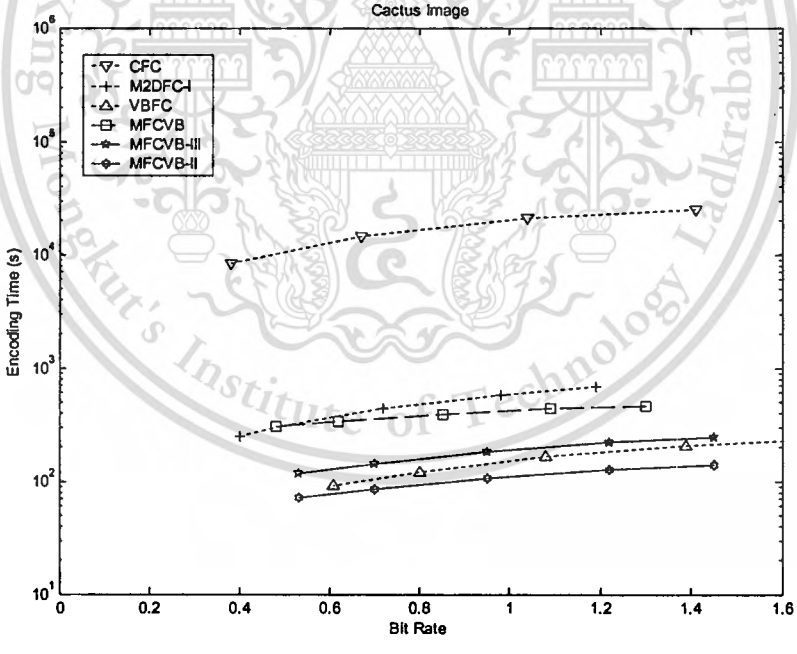
Figure 5.7 Comparison of the reconstructed field-based images of CFC, M2DFC-I, VBFC, MFCBV-I, MFCBV-II and MFCBV-III codecs.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



(c)

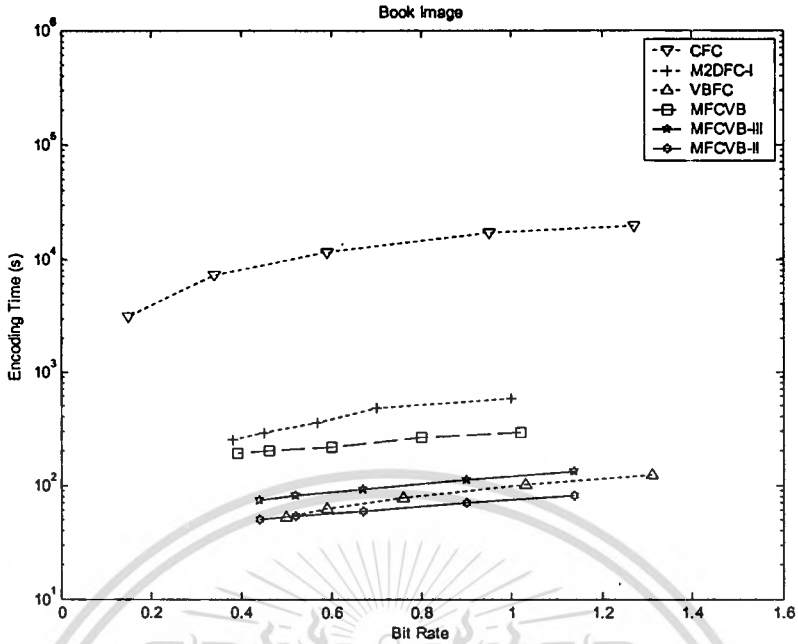
Figure 5.7 (Cont.)



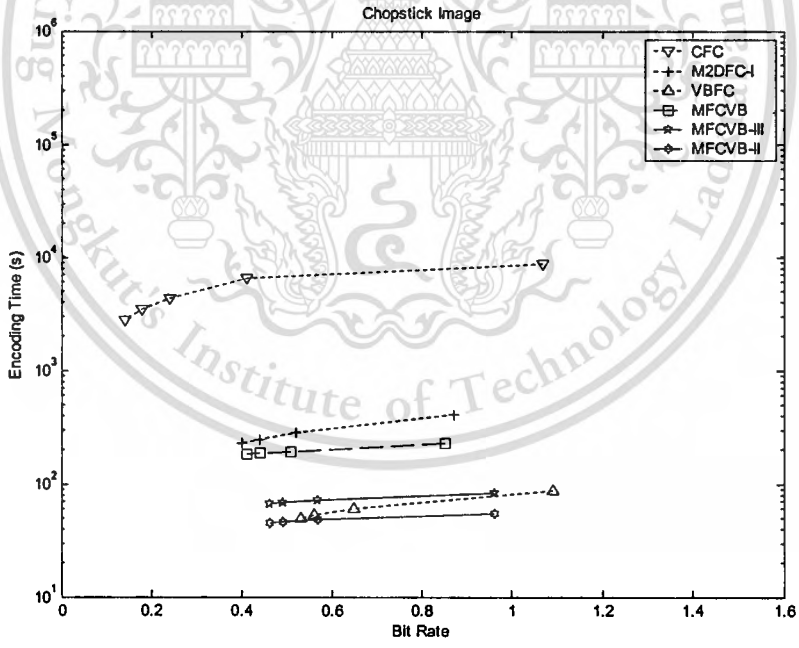
(a)

Figure 5.8 Comparison of encoding time of CFC, M2DFC-I, VBFC, MFCBV-I, MFCBV-II and MFCBV-III codecs.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



(b)



(c)

Figure 5.8 (Cont.)

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Table 5.6 Comparison of average encoding time of MFCBV-II and CFC.

Test image	MFCBV-II (s)	CFC (s)	Improved Ratio
Cactus	105.70	17,251.07	163
Book	62.47	11,676.23	187
Chopstick	39.01	5,137.30	132

Table 5.7 Comparison of average encoding time of MFCBV-III and CFC.

Test image	MFCBV-III (s)	CFC (s)	Improved Ratio
Cactus	181.79	17,251.07	95
Book	98.58	11,676.23	118
Chopstick	58.27	5,137.30	88

5.6 Conclusion

In this chapter, a novel fractal coding approach called *mutual fractal coding based on variance* is proposed. In this approach, the advantages of mutual fractal coding and state-of-the-art of variance-based fractal coding are combined together so that the reconstructed image quality, encoding time, decoding time and coding bit rate are improved. The mutual domain pool design is still the heart of the coding system, because it provides the following satisfactory conditions: (i) The mutual domain pools still preserved the similarity characteristic of images, retaining quality of reconstructed image as much as possible. (ii) The mutual domain pool size is smaller, thus decreasing the encoding time. (iii) The mutual fractal codes generated from the mutual domain-range mapping provide faster decoding. (iv) The mutual fractal codes require fewer bits, thereby reducing the coding bit rate. On the other hand, the variance-based approach is a simple and effective way to speed up the fractal encoding process, therefore in order to reach the objectives, the mutual domain pools are implemented for a variance-based scheme. The experimental results illustrate that overall efficiency of the proposed method can be summarized as follows. (i) The image quality of the proposed scheme is comparable with the CFC codec. (ii) The encoding time is greatly improved when compared with the CFC codec, but it is

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

slower than the VBFC by approximately two times. (iii) The decoding process converges in fewer iterations to decode the reconstructed image when compared with the CFC and VBFC codecs. It is approximately two times better. (iv) The coding bit rate is also reduced. Although successful effectiveness of the proposed coding system has been achieved, the computing of sorted mutual domain pools has been traded off. For this reason, the proposed method has still been comparable to the VBFC approach.

Finally, it is believed that the mutual domain pool design is just a beginning and it can be developed for improving the efficiency of the fractal coding.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Chapter 6

Conclusions and Suggestions

In this dissertation, two collaborating systems, acquisition and compression, have been investigated for reducing the redundancy of stereoscopic image and video sequences. The research work has been summarized as follows. In Chapter 2, an effective stereoscopic image and video acquisition system is proposed. This chapter first introduces the limitations of the conventional stereoscopic video acquisition system using double camera geometry, then its problems are formulated. These problems are (i) the difficulties of setting up the stereoscopic acquisition system and (ii) the widening image and video signals of a dual camera system. The proposed adapter for generating stereoscopic video is then analyzed and it is shown that the dual camera geometry and the intended geometry are comparable. Hence, the dual camera system can be completely replaced with the intended system. Finally, an archetype acquisition system was invented and experimented on to capture stereoscopic scenes. These scenes have come to be a set of stereoscopic test images. In this system, two controlling factors have been satisfied. The former is the elimination of complexity of stereoscopic acquisition system and the latter is the reduction of video signals.

Although a novel stereoscopic acquisition system has successfully invented and has satisfied the objectives, some limitations still remain. For instance, the system cannot be used to magnify image sequences when taking objects that are far away and the resolution of images is cut down by half, because of the characteristics of the field-sequential standard. In the case of stereoscopic video quality, it is dependent on the performance of the CCD camera. Finally, it is believed that the intended stereoscopic acquisition geometry can be replaced with high quality devices for improving the efficiency of system.

The rest of the chapters are concerned with coding systems. In Chapters 3 and 4, the novel concept of fractal coding is proposed to compromise the following objectives: (i) overwhelming the long encoding time, (ii) speeding up decoding time, (iii) retaining quality image and (iv) reaching a very-low-bit-rate condition. In order to achieve these objects, analyzing the construction and characteristics of field-

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

sequential stereoscopic image sequences was done. The problem was then formulated. In this phase, field-sequential stereoscopic video sequences differ from the typical video sequences in that they contain large amounts of high-vertical frequencies. These high frequencies are substantial for depth perception, therefore this characteristic is very important for designing a coding scheme. In order to overwhelm this obstacle, two fractal-coding approaches were proposed. The former is mutual scanline fractal coding (MSFC) for frame-based images and the latter is mutual 2-D fractal coding (M2DFC) for field-based images. The heart of these approaches is the mutual domain pool design. The mutual domain pools based on *equivalent* property of two complete metric spaces are designed for speeding up both encoding and decoding time. This method provides smaller domain pools, thus reducing the encoding time. The mutual fractal codes generated by this method also take fewer iterations for decoding the reconstructed images. This implies that the decoding time is shorter. In Chapter 4, the extended concept of mutual domain pool design to a two-dimensional approach is proposed. This extension makes the mutual fractal coding effective in two cases. The first is the reduction of coding bit rate and the other is the generic structure of the coding image.

In Chapter 5, mutual fractal coding based on variance is proposed. This method is a combination of the advantages of mutual fractal coding and variance-based fractal coding. The variance-based scheme is a simple and effective approach to improve fast encoding time and achieve good image quality, but the decoding time and coding bit rate are still no worse than the conventional fractal coding. On the other hand, although the mutual fractal coding can provide both less decoding time and a lower bit rate, the encoding time still has to speed up, therefore the combination of these approaches is an alternative way to improve the coding efficiency. Based on the experimental results, the following factors have been satisfactorily compromised: reconstructed image quality, encoding time, decoding time and coding bit rate.

Although the mutual domain pool design can greatly improve encoding and decoding time, some limitations still remain. For example, in the case of a larger image, the encoding and decoding time exponentially increases with the image size when using a simple searching technique, i.e., first-match searching technique. In this case, an extended version of the mutual domain pool design is being studied. This is the ***multiple domain pool design***. Finally, it is believed that the mutual domain pools and their extended concepts are also able to collaborate with the other techniques—

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

such as classifications [39], [46] feature vectors [46], frequency-like domains [47], [48], and others—for noticeably improving the fractal encoding and decoding time and image quality. In addition, hybrid approaches [49], [50] are also possible ways to cooperate with the approach taken in this dissertation.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

References

[1] T. Naemura, M. Kaneko, H. Harashima. "Compression and Representation of 3-D Images." **IEICE Trans. Inf. & Syst.**, vol. E82-D, no. 3, 1999, pp. 558-567.

[2] M. Okui, A. Hanazato, F. Okano, I. Yuyama. "A study on Scanning Methods for a Field-Sequential Stereoscopic Display." **IEEE Trans. Circuits Syst. Video Technol.**, vol. 10, no. 2, 2000. pp. 244-260.

[3] M.G. Perkins. "Data Compression of Stereo pairs." **IEEE Trans. on Commun.**, vol. 40 no. 4, April 1992, pp. 684-696.

[4] J. Konrad. "Visual Communications of Tomorrow: Natural, Efficient, and Flexible." **IEEE Comm. Mag.** Jan. 2001, pp 126-133.

[5] J. Mulligan, X. Zabulis, N.Kelshikar, K.Daniilidis. "Stereo-Based Environment Scanning for Immersive Telepresence." **IEEE Trans. Circuits Syst. Video Technol.**, vol. 14, no. 3, 2004. pp. 304-320.

[6] L.M.J. Meesters, W.A. Ijsselsteijn, P.J.H. Seuntiens. "A Survey of Perceptual Evaluations and Requirements of Three-Dimensional TV." **IEEE Trans. Circuits Syst. Video Technol.**, vol. 14, no. 3, 2004. pp. 381-391.

[7] F. Isgro, E. Trucco, P. Kauff, O. Schreer. "Three-Dimensional Image Processing in the Future of Immersive Media." **IEEE Trans. Circuits Syst. Video Technol.**, vol. 14, no. 3, 2004. pp. 288-303.

[8] A.Woods. "Stereoscopic Presentations–Taking the Difficulty out of 3D." **Proc. 6th International Workshop on 3-D Imaging Media Technology**, pp. 1-6, Seoul, Korea. Jul. 2000.

[9] H. Huang, C. Kao, Y. Hung. "Generation of Multiviewpoint Video from Stereoscopic Video." **IEEE Trans. Consumer Electronics**, vol. 45, no. 1, 1999. pp. 124-134.

[10] H. Mitsumine, Y. Yamanouchi, S. Inoue. "An Acquisition Method of 3-Dimensional Video Components for Image-Based Virtual Studio." **Proc. IEEE International Conference on Image Processing**, pp.1101-1104, Oct. 2001.

[11] L. Tai, R. Jain. "3D Video Generation with Multiple Perspective Camera Views." **Proc. IEEE International Conference on Image Processing**, pp. 9-12, Oct. 1997.

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

[12] A. Woods, T. Docherty, R. Koch. "3D Video Standards Conversion." **Stereoscopic Displays and Applications VII, Proc. SPIE,** San Jose, California, Feb. 1996.

[13] A. Woods, T. Docherty, R. Koch. "Image Distortions in Stereoscopic Video Systems." **Stereoscopic Displays and Virtual Reality Systems VI, Proc. SPIE,** pp. 36-48, San Jose, California, 1993.

[14] B. Wohlberg, G. de Jager. "A Review of the Fractal Image Coding Literature." **IEEE Trans. Image Processing,** vol. 8, no. 12, 1999. pp. 1716-1729.

[15] P. Limmaneepraserth, R. Varakulsiripunth. "Designing and Implementing Mutual Scan-line Fractal Coding for Field-sequential Stereo Video Sequences." **Proc. 7th International Symposium on Signal Processing and Its Applications,** pp. 109-112, Paris, France, Jul. 2003.

[16] R.D. Kell. "Television System." **US Patent 2508920** 23rd May 1950.

[17] J.F. Butterfield. "Stereoscopic Television System." **US Patent 4734756** 29th Mar. 1988.

[18] J.A. Roese. "PLZT Stereoscopic Television System." **US Patent 3903358** 2nd Sept. 1975.

[19] D.J. Montgomery, C.K. Jones, J.N. Stewart, A. Smith. "Stereoscopic Camera Design." **Stereoscopic Displays and Virtual Reality Systems IX, Proc. SPIE,** pp. 26-37, San Jose, California, Jan. 2002.

[20] P.A. Femano. "Stereoscopic Conversion Assembly for Closed Circuit 2-D Television System." **US Patent 4943852** 24th Jul. 1990.

[21] H. Sudo. "Stereoscopic Television System." **US Patent 5003385** 26th Mar. 1991.

[22] H. Yamaguchi, Y. Tatehira, K. Akiyama, Y. Kobayashi. "Stereoscopic Images Disparity for Predictive Coding." **Proc. International Conference on Acoustics, Speech, and Signal Processing,** pp. 1976-1979 May 1989.

[23] M.S. Moellenhoff, M.W. Maier. "DCT Transform Coding of Stereo Images for Multimedia Applications." **IEEE Trans. Industr. Electron.,** vol. 45, no. 1, 1998. pp. 38-43.

[24] M.S. Moellenhoff, M.W. Maier. "Transform Coding of Stereo Image Residuals." **IEEE Trans. Image Processing,** vol. 7, no. 6, 1998. pp. 804-812.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

[25] T. Frajka, K. Zeger. "Residual Image Coding for Stereo Image Compression." **Proc. IEEE International Conference on Image Processing**, pp. 217-220, Sept. 2002.

[26] H. Aydinoglu, M.H. Hayes. "Stereo Image Coding: A Projection Approach." **IEEE Trans. Image Processing**, vol. 7, no. 4, 1998. pp. 506-516.

[27] M.Y. Nayan, E.A. Edirisinghe, H.E. Bez. "Baseline JPEG-Like DWT CODEC for Disparity Compensated Residual Coding of Stereo Images." **Proc. 20th Eurographics UK Conference**, pp. 67-74 Jun. 2002.

[28] A. Bovik. **Handbook of Image and Video Processing**, Academic Press, 2000.

[29] M. Johanson. "Stereoscopic Video Transmission over the Internet." **Proc. 2nd IEEE Workshop on Internet Applications**, pp. 12-19, Jul. 2001

[30] M.W. Burke. **Image Acquisition: Handbook of Machine Vision Engineering.** Volume I, Chapman & Hall, 2-6 Boundary Row, London SE1 8HN, UK, 1996.

[31] J. Gluckman, S.K. Nayar. "Planar Catadioptric Stereo: Geometry and Caribration." **Proc. IEEE Conference on Computer Vision and Pattern Recognition**, 1999.

[32] R.C. Gonzales, Richard E. Woods. **Digital Image Processing**, Addison-Wesley Publishing Company, 1992.

[33] Y. Miyakawa et al.. "Synchronized Three Dimensional Imaging Apparatus." **US Patent 5028994** 2nd Jul. 1991.

[34] A. E. Jacquin. "Image Coding Based on a Fractal Theory of Iterated Contractive Image Transformations." **IEEE Trans. Image Processing**, vol. 1, no. 1, 1992. pp. 18-30

[35] H. T. Chang, C. J. Kuo. "Iteration-Free Fractal Image Coding Based on Efficient Domain Pool Design." **IEEE Trans. Image Processing**, vol. 9, no. 4, 2000. pp. 329-339.

[36] A. J. Woods, D. Offszanka, G. Martin. "A PC-based stereoscopic video walkthrough." **Stereoscopic Displays and Virtual Reality Systems VI, Proc. SPIE**, pp. 306-312, San Jose, California, 1999.

[37] M. F. Barnsley. **Fractals Everywhere Second Edition**. Academic Press Professional, London, 1993.

[38] J. Kominek. "Advances in Fractal Compression for Multimedia Applications." **Multimedia Systems**, vol. 5, no. 4, 1997. pp. 255-270.

This material is protected by copyright. Use only for non-commercial use.
Forbidden to modify the content, and cite the document when use.

[39] Y. Fisher. **Fractal Image Compression: Theory and Application**. Springer-Verlag, New York, 1995.

[40] J. Kominek. "Convergence of Fractal Encoded Images." **Proc. DDC**, Snowbird, UT, USA, 1995. pp. 242-251.

[41] A.E. Jacquin. "Fractal Image Coding: A Review." **Proc. IEEE**, vol. 81 no. 10, 1993. pp. 1451-1465.

[42] D. Saupe, R. Hamzaoui. "Complexity Reduction Methods for Fractal Image Compression." **Proc. International Conference on Image Processing; Mathematical Methods and Applications**, Sept. 1994.

[43] R.L. Kruse, B.P. Leung, C.L. Tondo. **Data Structures and Program Design in C**, Prentice-Hall International Editions, 1991.

[44] C.K. Lee, W.K. Lee. "Fast Fractal Image Block Coding Based on Local Variances." **IEEE Trans. Image Processing**, vol. 7, no. 6, 1998. pp. 888-891.

[45] C. He, S.X. Yang, X. Huang. "Variance-Based Accelerating Scheme for Fractal Image Encoding." **Electron. Lett.**, 2004, 40, (2), pp. 115-116.

[46] M. Polvere, M. Nappi. "Speed-up in Fractal Image Coding: Comparison of Methods." **IEEE Trans. Image Processing**, vol. 9, no. 6, 2000. pp. 1002-1009.

[47] T.K. Truong, J.H. Jeng, I.S. Reed, P.C. Lee, A.Q. Li. "A Fast Encoding Algorithm for Fractal Image Compression Using the DCT Inner Product." **IEEE Trans. Image Processing**, vol. 9, no. 4, 2000. pp. 529-535.

[48] J.H. Jeng, T.K. Truong, J.R. Sheu. "Fast Fractal Image Compression Using the Hadamard Transform." **IEE Proc.-Vis. Image Signal Process.**, vol. 147, no. 6, 2000.

[49] C.S. Tong, M. Pi. "Analysis of a Hybrid Fractal-Predictive-Coding Compression Scheme." **Signal Processing: Image Communication**, vol. 18, 2003, pp. 483-495.

[50] Z. Wang, D. Zhang, Y. Yu. "Hybrid Image Coding Based on Partial Fractal Mapping." **Signal Processing: Image Communication**, vol. 15, 2000, pp. 767-779.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Appendix

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Journal of Signal Processing, Vol. 9, No. 2, pp. 165-178, March 2005

PAPER

Mutual Fractal Coding for Field-Sequential Stereoscopic Video

Ruttikorn Varakulsiripunth[1] and Para Woraratpanya[2]

Faculty of Engineering and Research Center for Communications and Information Technology
King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand
E-mail: kvruttik@kmitl.ac.th[1] and para@kmitnb.ac.th[2]

Journal of Signal Processing

信号処理

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Journal of Signal Processing, Vol. 9, No. 2, pp. 165–178, March 2005

PAPER

Mutual Fractal Coding for Field-Sequential Stereoscopic Video

Ruttikorn Varakulsiripunth[1] and Para Woraratpanya[2]

Faculty of Engineering and Research Center for Communications and Information Technology
King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand
E-mail: kvruttik@kmitl.ac.th[1] and para@kmitnb.ac.th[2]

Abstract **Fractal coding is a recent compression technique based on self-similarity of images. This technique has various advantages in terms of resolution independence, fast decoding, and high compression ratio. However, in the encoding phase it takes a long computational time. In order to overcome this problem, we propose a novel coding approach, i.e., mutual fractal coding. In this approach, there is a significant characteristic that differs from the conventional fractal coding. That is the effective domain pool design; the mutual domain pools based on *equivalent* property of two complete metric spaces are designed for speeding up the encoding and decoding time. For applications to field-sequential stereoscopic video sequences, the remaining problem is that the high vertical-frequencies of stereoscopic image pair in frame-based images are suppressed. This makes the depth perception defective. In this case, there are two simple techniques for overcoming this problem. The first technique is mutual scanline fractal coding (MSFC) for frame-based images. The other is mutual 2-D fractal coding (M2DFC) for field-based images. The experimental results illustrate that the intended approaches can reduce both encoding and decoding times when compared with the conventional fractal coding based on quadtree partition. Moreover, the reconstructed image quality is improved in terms of preserving the high vertical-frequencies by using the MSFC for frame-based images and M2DFC for field-based images.**

Keywords: fractal coding, mutual scanline fractal coding, mutual 2-D fractal coding, self-similarity, stereoscopic image pair, stereoscopic video sequence

1. Introduction

Currently, stereoscopic video is becoming an important ingredient of multimedia applications [1], [2]. Especially, the field-sequential stereoscopic video is one of the most practical approaches for producing and displaying stereoscopic image sequences on existing equipment [2]. Nevertheless, it requires an effective coding to minimize its information for storage spaces and transmission bandwidths. Thus, in this paper the effective coding technique and the characteristics of the field-sequential stereoscopic video are collaboratively investigated.

In general, the field-sequential stereoscopic video sequences are composed of left and right images acquired from two slightly different viewpoints, thus making them similar and containing a lot of redundant information. In order to remove this redundancy, an effective coding technique is essential. With a characteristic of the field-sequential stereoscopic video sequences containing a lot of self-similarity information, the suitable coding approach for such information is fractal coding based on the self-similarity property. Furthermore, this technique has various advantages in terms of resolution independence, fast decoding, and high compression ratio. However, its long encoding time is still a main drawback. Moreover, in case of frame-based image coding of the field-sequential stereoscopic video sequences, the conventional 2-D fractal coding technique makes the depth or three-dimensional perception defective, since the high vertical-frequencies of a stereoscopic image pair are suppressed.

In order to conquer these problems, a novel coding approach, mutual fractal coding, for coding the field-sequential stereoscopic video sequences is proposed. In this approach, there is a significant characteristic that differs from the design of the conventional fractal coding; that is the effective domain pool design. It is one of the important phases for speeding up the encoding time of fractal coding. A. Jacquin [3] concluded his unsolved problems concerning with the conception of fractal encoding. The domain pool is one of his lists. Since then, much less paper [4] has been proposed on the efficient domain pool design, but such paper has been emphasized to the iteration-free for decoding process. B. Wohlberg [5] discussed the survey literatures related to domain pool design and concluded that the simplest way to speed up the fractal encoding is domain pool reduction. However, this method is a trade-off between domain pool size and image quality; the smaller domain pool leads to the poorer image quality. Therefore, the objectives of this paper are to design the effective domain pool whose size is reduced as small as possible and to retain the image quality as good as feasible. For this reason, we have proposed a novel approach for designing the domain pool, i.e., mutual domain pool design. The mutual domain pools based on *equivalent* property of two complete metric spaces are designed for speeding up the encoding and decoding time. This method provides the smaller domain pools, thus reducing the searching time of domain-range mapping. The codes generated by these mutual domain pools are called *mutual fractal codes*.

The remaining problem is that the high vertical-frequencies of the stereoscopic image pair are suppressed. In this case, there are two methods to avoid this problem. The former is mutual scanline fractal coding for frame-based images. In this method, the scanline technique is used to partition the range block of images. It is a simple and practical technique to preserve the high vertical-frequencies of the stereoscopic image pairs. Furthermore, it also avoids the image-crosstalk problem caused by blending the left and right images [6]. The latter is mutual 2-D fractal coding (M2DFC) for field-

This material is reserved for educational use only, not allowed for commercial use. Forbidden to modify the content, and cite the document when use.

based images. In this approach, the de-interleaved technique is utilized for eliminating the high vertical-frequencies of the stereoscopic image pairs. This technique is employed in the preprocessing step of the M2DFC.

The rest of this paper is organized as follows: Section 2 reviews the construction of a field-sequential stereoscopic video sequence and its characteristics. The high vertical-frequencies of stereoscopic image pairs are also addressed. In Section 3, the theoretical model of mutual fractal coding is formed. Then, we propose the design of coding system in Section 4, In Section 5, the implementations and applications of the intended approach performed in both cases: frame- and field-based coding are described. The experimental results are illustrated in Section 6. Finally, the conclusions are given in Section 7.

2. A Brief Review of Field-Sequential Stereoscopic Video Sequences

In this section, we concisely describe the construction of a field-sequential stereoscopic video sequence and its characteristics.

2.1 Construction of field-sequential stereoscopic video sequences

Generally, the field-sequential stereoscopic video is made up of a sequence of stereoscopic image pairs. Fig. 1 shows its construction. This construction is compatible with the typical scanning of television systems for which every frame consists of two fields: even and odd fields. The number of times for repainting the screen in one second is called *field rate*. This field rate is customarily related to a vertical synchronous frequency, 60 Hz for NTSC system, for instance. Hence, with this field rate, it is so expeditious that the human eyes cannot identify the alternation between the even and odd fields. Thus, the field-sequential scanning is an appropriate technique for occupying the left and right images in one frame, consisting of the even field for the left image and the odd field for the right image, and for displaying on standard televisions and computer monitors. A model of the field-sequential stereoscopic video acquisition system proposed in [7] is illustrated in Fig. 2. Fig. 3(a) shows an example of a stereoscopic image pair acquired by this model.

2.2 Characteristics of field-sequential stereoscopic image pairs

Fundamentally, an image can be viewed as a two-dimensional signal composed of horizontal and vertical axes that denote horizontal and vertical frequencies, respectively. For a field-sequential stereoscopic image, both horizontal and vertical frequencies have great difference as illustrated in Fig. 4, but for a typical image, they have no great difference as graphically depicted in Fig. 5. At this point, it is summarized that the field-sequential stereoscopic video sequence differs from the typical video sequence in that it contains a large amount of high frequencies in vertical direction of each frame. These high frequencies are substantial for depth perception. Hence, this characteristic of the field-sequential stereoscopic video sequence is very important for designing

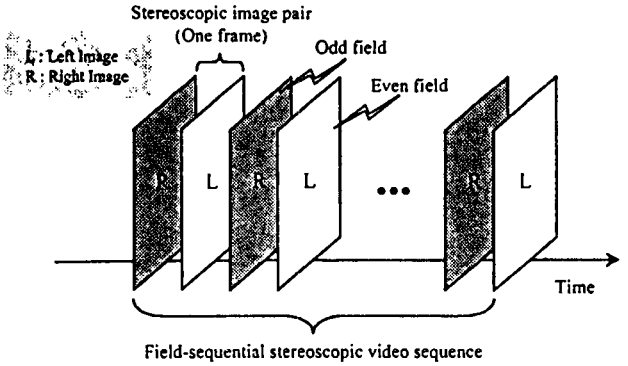


Fig. 1 Construction of field-sequential stereoscopic video

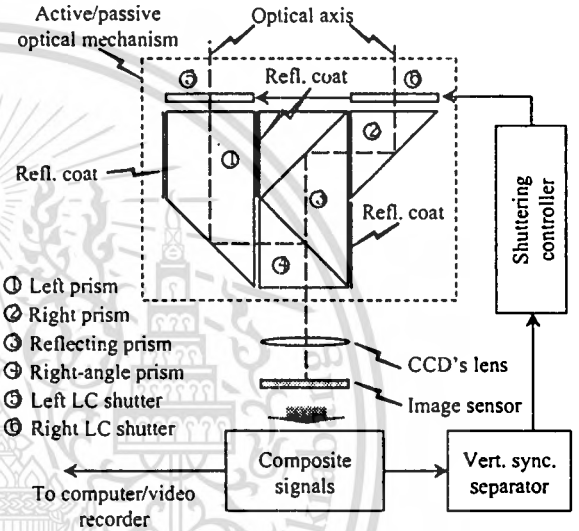


Fig. 2 Model of stereoscopic video acquisition system using active/passive optical mechanism

a coding scheme described in the following sections. However, in order to preserve the high vertical-frequencies, we can apply the MSFC to the frame-based image depicted in Fig. 3(a) and the M2DFC to the field-based image shown in Fig. 3(b). These approaches are designed and implemented with the theoretical model described in the next section.

3. Theoretical Model

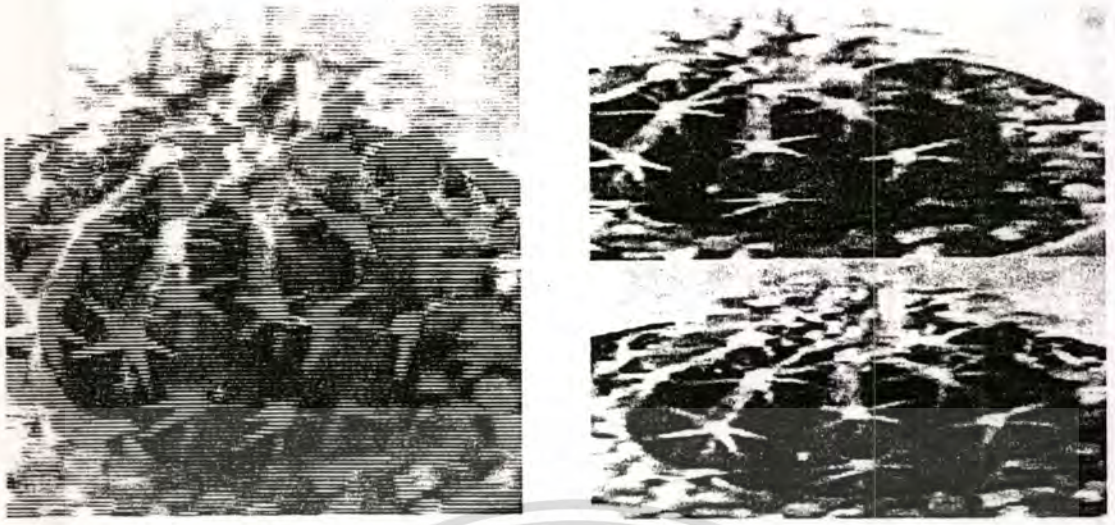
In this section, theoretical model of mutual fractal coding is concisely presented. This model is made of the following principal theories: metric spaces, affine transformations, contractive mapping fixed-point theorem, and partitioned iterated function system (PIFS). However, only the relevant definitions and theorems are described for lack of space. For additional details of above theories, the reader should refer to [8]. In addition, we describe solely the theory of scanline model. For the extended (2-D) model, the implementation scheme is illustrated in Subsection 5.2.

3.1 Model of mutual fractal codes

We begin with describing the definition of elementary mathematical model for the mutual scanline fractal coding,

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and



(a) (b)

Fig. 3 (a) Original 256 x 256 pixels, 8 bpp "cactus" stereoscopic image pair in field-sequential format (frame-based image) and (b) field-based image separated from the original stereoscopic image pair

that is, *complete* metric spaces of a stereoscopic image pair and their *equivalent* property. Typically, the fractal image coding requires such metric spaces to be complete in order that they have no missing values. Furthermore, the equivalent property of two complete metric spaces is also important for designing the mutual domain pools. The significant condition of two equivalent metric spaces is that there is no extreme stretching or compression of the spaces [8]. With this condition, the stereoscopic image pair can be viewed as a mutual deformation of images. That is, a left image of the stereoscopic image pair is a deformation of a right image, and vice versa. In practice, the deformation mechanisms of a stereoscopic video acquisition system are the left and right prisms as schematically shown in Fig. 2. For this reason, let (X,d) be a complete metric space of the stereoscopic image pair. Then it can be decomposed into two complete metric spaces (X_L,d_L) and (X_R,d_R). The equivalence of two complete metric spaces is defined as follows:

Definition 1: *Let (X_L,d_L) and (X_R,d_R) be complete metric spaces of the left and right discrete images of a stereoscopic image pair. Then, these complete metric spaces are equivalent if there is a function $f: X_L \to X_R$ that is one-to-one and onto, such that the metric $\tilde{d}_L$ on X_L defined by*

$$\tilde{d}_L(x,y) = d_R(f(x),f(y)),\ \forall x,y \in X_L$$

is equivalent to d_L. □

After introducing the definition of metric spaces with their necessary properties, then transformations on these metric spaces are addressed. In case of the mutual scanline fractal coding, the transformations used to generate the mutual fractal codes are a pair of affine transformations and are contractive mappings [8]. The condition that makes these transformations contractive is defined as follows:

Definition 2: *Let (X_L,d_L) and (X_R,d_R) be complete metric spaces and let $f: X_L \to X_R$ and $g: X_R \to X_L$ be affine transformations. These transformations f and g are called contractive mappings if there exist constants $0 \le s_L, s_R < 1$ such that*

$$d_R(f(x),f(y)) \le s_L \cdot d_L(x,y),\ \forall x,y \in X_L$$

and

$$d_L(g(x),g(y)) \le s_R \cdot d_R(x,y),\ \forall x,y \in X_R$$

Such constants s_L and s_R are called contractivity *factors for f and g, respectively.* □

Theorem 1: *From Definition 2, the contractivity factor of $(f \circ g)(x) = f(g(x))$, where $f \circ g$ is the composition of f and g, is $s_L \cdot s_R$.*

Proof: See Appendix. □

Definitions 1 and 2 provide the suitable complete metric spaces and contractive transformations, respectively, for generating mutual fractal codes. Based on these definitions, a pair of contractive transformations f and g is imposed to mutually map between two complete metric spaces (X_L,d_L) and (X_R,d_R). Moreover, Theorem 1 also proves that $(f \circ g)(x)$ is contractive and its contractivity factor is $s_L s_R$. However, in order to solve the *image encoding problem* (or *inverse problem*) [3], [9], the fixed-point theorem is very important. Accordingly, based on Definitions 1 and 2, and Theorem 1, the contractive mapping fixed-point theorem presented in [3], [9] can be extended to our objective as Theorem 2.

Theorem 2: *Let (X_L,d_L) and (X_R,d_R) be equivalent complete metric spaces, and let $f: X_L \to X_R$ and $g: X_R \to X_L$ be contractive mappings. Then $f \circ g$ possesses exactly one fixed point $x_f \in X$, such that $X = X_L \cup X_R$ and $X_L \cap X_R = \varnothing$. Furthermore, for any point $x \in X$, the sequence*

$$\{(f \circ g)^{\circ n}(x) : n = 0, 1, 2, \ldots\}$$

converges to x_f. That is,

$$x_f = (f \circ g)(x_f) = \lim_{n \to \infty} (f \circ g)^{\circ n}(x), \qquad \text{for each } x \in X$$

Proof: See Appendix. □

This material is reserved for educational use only. Illegal use for commercial use.
Please read the copyright statement, and cite the document when use.

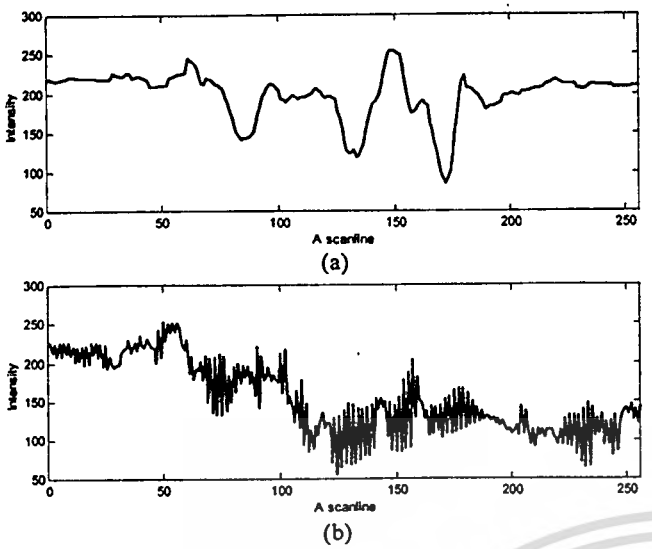


Fig. 4 The representation of (a) the horizontal image scanline in row 51st and (b) vertical image scanline in column 120th of the original stereoscopic image pair in Fig. 3(a)

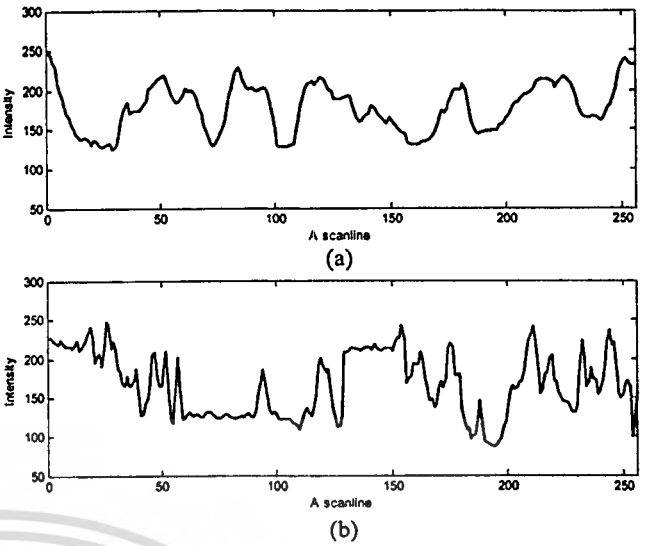


Fig. 5 The representation of (a) the horizontal image scanline in row 51st and (b) vertical image scanline in column 60th of the field-based image in Fig. 3(b)

Theorem 2 proves that $f \circ g$ still possesses exactly one fixed point, even though it performs on two equivalent complete metric spaces. In practice, this provides us an effective design of mutual domain pools in Subsection 4.2. Now, Theorem 2 is promptly applied to a *partitioned iteration function system* (PIFS). Nevertheless, in this paper, we also make use of divide-and-conquer technique to provide suitable conditions for PIFS. This technique involves solving a particular computational problem by *dividing* it into more subproblems of smaller size, and then *merging* the solutions of subproblems to produce a solution of the original problem. Therefore, for scanline case, we firstly divide an image into a number of pairs of scanlines. Then, each pair of scanlines is defined as follows: D_L and D_R denote the mutual domain pools of scanlines of left and right images, respectively, whereas R_L and R_R denote the scanlines of range blocks. Finally, the following definition generates the mutual fractal codes.

Definition 3: *Let (X_L, d_L) and (X_R, d_R) be complete metric spaces of the left and right discrete images of a stereoscopic image pair, and let $D_L \subset X_L$ and $D_R \subset X_R$. Then a* partitioned iteration function system (PIFS) *is a collection of contractive mappings:*

$$w_u : D_{L(u)} \to R_R \text{ for } u = 1,2,\ldots,N$$

and

$$w_v : D_{R(v)} \to R_L \text{ for } v = 1,2,\ldots,M$$

where N and M are a number of domain blocks of D_L and D_R, respectively. □

Definition 3 makes fractal image encoding practical. Therefore, the mutual fractal codes for a whole image can be generated by means of the following definition.

Definition 4: *By Definition 3, the mutual fractal codes W of an encoded image can be obtained as*

$$W = \bigcup_{k=1}^{K/2}\left(\left(\bigcup_{u=1}^{N} w_{u,k}\right) \cup \left(\bigcup_{v=1}^{M} w_{v,k+1}\right)\right)$$

where N and M are a number of range blocks of R_L and R_R, respectively, and K is a total scanline of a stereoscopic image pair. □

Now, we can make the theory more practical for generating the mutual fractal codes based on the previous definitions and theorems. The following subsections describe the principle of mutual scanline fractal encoding and decoding.

3.2 Mutual scanline fractal encoding

Encoding is a step to encode the mutual fractal codes generated by Definitions 1 to 4 and Theorem 2. The heart of this phase is to determine the optimal values of parameters of domain-range mappings. To account for manipulation of the mutual scanline fractal encoding, there are two issues that are introduced. Firstly, we assume that there are mutual do-main pools D_L and D_R, which denote the scanlines of a stereoscopic image pair, and their intersection ($D_L \cap D_R = \emptyset$) is empty. This implies that the mutual domain pools are not unified in spatial coordinates. Moreover, the mutual domain pools satisfy Definition 1. Secondly, there is a pair of affine transformations $f : D_L \to R_R$ and $g : D_R \to R_L$, where R_R and R_L are the scanlines of range blocks, and these transformations satisfy Definition 2.

In practice, the process of generating the mutual fractal codes based on Definition 3 requires two types of data blocks, namely, domain and range blocks. Usually, the size of the domain blocks is greater than that of the range blocks. For establishing the range blocks, each scanline of original image is partitioned into a set of nonoverlapping blocks. Sequentially, in order to form mutual domain pools, where the domain blocks belong, each scanline of original image is partitioned into a set of overlapping blocks. Thus, the mutual domain pools are a collection of overlapping blocks. After providing the appropriate domain and range blocks, it reaches domain-range mapping process. In this process, there are two main steps for mutually obtaining the best matching between domain and range blocks. As depicted in Fig. 6(a), the

This material is received for educational use. Forbidden to modify the content, and distribute document without use.

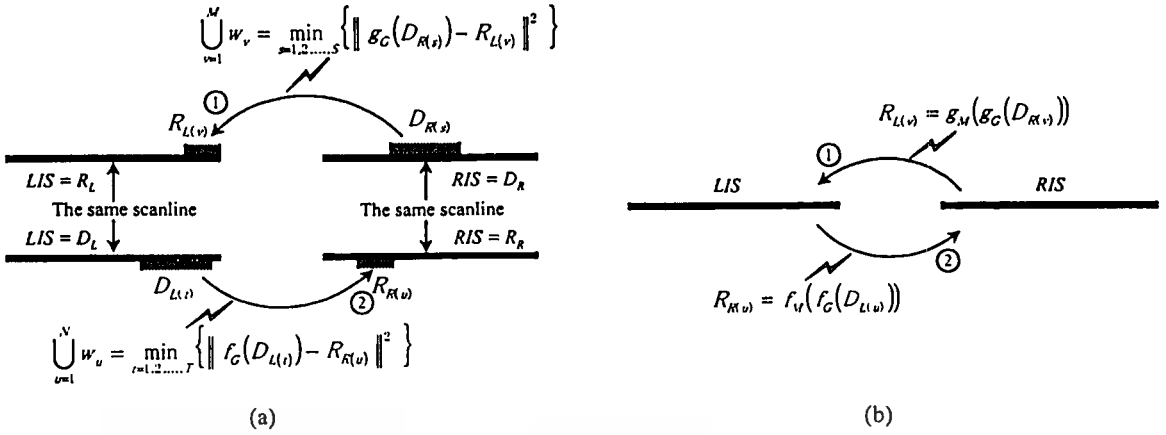


Fig. 6 (a) Encoding and (b) decoding steps of mutual fractal codes

mutual domain pools consist of D_L and D_R. Each range block of R_L and R_R mutually searches for the best matching of a domain block in mutual domain pools D_R and D_L, respectively. That is, in the first step, each range block of the scanline $R_{L(v)}$ searches for the best matching of the domain blocks in the domain pool $D_{R(s)}$ by means of

$$\bigcup_{v=1}^{M} w_v = \min_{s=1,2,\ldots,S} \left\{ \left\| g_G(D_{R(s)}) - R_{L(v)} \right\|^2 \right\} \quad (1)$$

where S is a total possible domain block in the domain pool D_R, and $g_G(.)$ denotes the geometric transformation described in the end of this subsection. Then, in the second step, each range block $R_{R(u)}$ searches for the best matching of a domain block in the domain pool $D_{L(t)}$ by means of

$$\bigcup_{u=1}^{N} w_u = \min_{t=1,2,\ldots,T} \left\{ \left\| f_G(D_{L(t)}) - R_{R(u)} \right\|^2 \right\} \quad (2)$$

where T is a total possible domain block in the domain pool D_L, and $f_G(.)$ denotes the geometric transformation.

The tool used to measure the similarity of the domain and range blocks is composed of geometric and massic transformations. In Eqs.(1) and (2), geometric transformations are $g_G(.)$ and $f_G(.)$ whereas massic transformations are implicit in the form of least-square error (Subsection 4.4 describes more details in this issue.) The step of generating mutual fractal codes W is that the domain blocks are transformed by using contractive affine (geometric) transformations. Then, the root-mean-square (rms) metric (massic transformation) is applied to measure the similarity of the domain and range blocks. Finally, the mutual fractal codes generated by these transformations are uniformly quantized so as to be the compact forms.

3.3 Mutual scanline fractal decoding

The concept of decoding the mutual fractal codes is very simple. It involves the obtaining of a fixed point of a contractive transformation. That is, the decoding process is an iteration of the contractive transformation on any initial values till reaching an approximate fixed point. However, the mutual scanline fractal decoding differs from the conventional one. In addition, the concept of decoding is based on the fact that if the seed image is initialized close to the original image, then the contractive transformation converges rapidly to a fixed point [10]. Therefore, in order to describe the manipulation of this decoding process, the left image scanline (*LIS*) and right image scanline (*RIS*) are set up with zero-initial values, i.e., all pixels of the scanlines are zeroes. Here, the seed image scanlines are *LIS* and *RIS*. At each iteration, there are two steps to decode the mutual fractal codes as depicted in Fig. 6(b). In the first step, all range blocks of the *LIS* are mapped with the appropriate domain-transformed blocks from the *RIS*, such that $R_{L(v)}=g_M(g_G(D_{R(v)}))$, where $g_G(.)$ and $g_M(.)$ are contractive (geometric) and massic transformations, and $D_{R(v)}$ is a domain block of the *RIS*. Then, in the second step, the outcome in the *LIS*, which is a domain-block scanline of the *RIS*, is performed with the same process as the first steps handle. Note that the *LIS*, a seed image scanline of *RIS*, is not initialized with zero values, but holds the results from the first step. This implies that the seed image scanline *LIS* is initialized close to the original one. Thus, the transformation $g_M(f_G(.))$ converges rapidly to the fixed point. This is an effective method for generating the mutual fractal codes that can speed up the decoding time.

4. Design of Coding System

In designing phase, there are five significant issues that have to be considered and decided for a coding scheme. These issues are scanline partition, mutual domain pool design, mutual scanline transformations, distortion measure, and bit allocation. In addition, all issues described in this section can be extended easily to 2-D case.

4.1 Scanline partition

Scanline partition is a procedure to divide the scanline of the original image into nonoverlapping one-dimensional (1-D) blocks. These blocks are called *range* blocks. The size of range block is an important factor that influences directly to bit rate. This means that the larger range blocks lead to the lower bit rate. On the other hand, the smaller size captures the more details of image scanlines.

Personal use is permitted, but republication/redistribution requires IEEE permission.
Please see http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

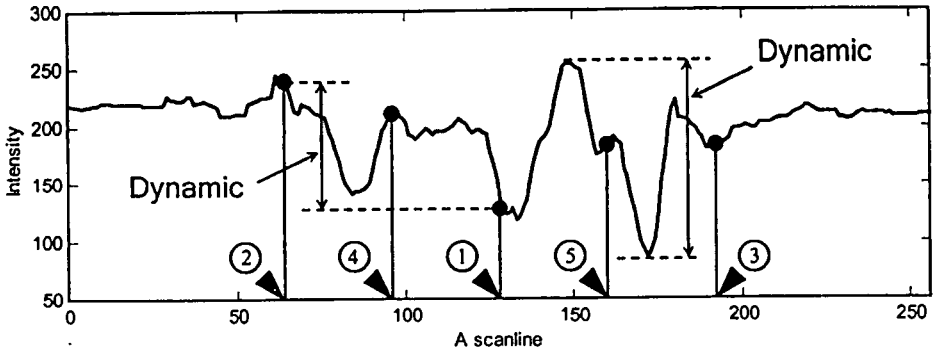


Fig. 7 Illustration of adaptive-size partition by using dynamic range, which refers to the extremes of intensity change of an image scanline in any intervals

In this paper, we propose two kinds of partitions: fixed- and adaptive-size partitions. For the former, each scanline is equally divided into a fixed size, such as 1x8, 1x16, or 1x32. This method is very simple, but not flexible for the bit rate. For the latter, the binary-tree partition technique is used to divide the image scanline. In this method, each scanline is equally divided into two subparts. Then, the *dynamic* range, which refers to the extremes of intensity change of an image scanline in any interval as depicted in Fig. 7, is used to decide splitting. If the dynamic range value of each subpart exceeds the threshold, then that subpart is equally divided. Otherwise, do nothing. This task is repeatedly processed in the previous fashion till reaching the restricted block size. In Fig. 7, the numbers in circle are indicated the sequence of partition. Hence, the dynamic range threshold is an important parameter to control a number of range blocks. For the lower bit rate and image quality, the smallest and largest sizes of bounded range blocks are defined, i.e., 1x8 and 1x64 are of the smallest size and the largest size, respectively. By this method, the adaptive partition makes the bit rate flexible, and also achieves better image quality.

4.2 Mutual domain pool design

A domain pool is a place where the range block searches for the best matching block. The good design of domain pool leads to the less encoding time of fractal coding [5]. In this paper, we propose a simple method for designing domain pools by utilizing the equivalent property of two complete metric spaces as described by Definition 1 in Subsection 3.1. That is, the coding scheme is designed with two domain pools, D_L and D_R. These domain pools are independent, but they are equivalent. In the searching step, each range block on the space R_L finds the best matching domain block in the space D_R, and vice versa. This technique makes the domain pool size smaller, thus reducing the encoding time.

4.3 Mutual scanline transformations

In this subsection, we propose a pair of scanline transformations designed for our codec (coder and decoder). These transformations are 1-D affine transforms defined as:

$$\begin{cases} f_G(x_l) = ax_l + b, & \forall x_l \in X_L \\ g_G(x_r) = cx_r + d, & \forall x_r \in X_R \end{cases} \quad (3)$$

where a and c are geometrical contractivity factors [3], [8], [12] of the transformations, b and d are translation factors, and x_l and x_r are range blocks on the complete metric spaces (X_L,d_L) and (X_R,d_R), respectively. In practice, a and c are set to a constant value, usually equal to 1/2, and b and d are integers. $f_G(x_l)$ and $g_G(x_r)$ are domain-transformed blocks on the complete metric spaces (X_L,d_L) and (X_R,d_R), respectively. The steps of generating mutual fractal codes are as follows. Firstly, each domain block whose size is greater than range block in the metric space (X_L,d_L) is transformed by f_G in order to find the best matching of range block in the metric space (X_R,d_R). At the same time, each domain block in the metric space (X_R,d_R) is transformed by g_G in order to find the best matching of range block in the metric space (X_L,d_L). In this case, the tool for measuring the best matching block is root-mean-square (rms) metric presented in the next subsection.

4.4 Distortion measure

Distortion measure is a tool to measure the similarity of two image scanlines. In practice, the tool commonly used for measuring the distortion of domain-range mapping is the root-mean-square (rms) metric. This metric is easy to compute the optimal values of parameters s_j and o_j, where s_j and o_j denote contrast and brightness, respectively, in Eq.(4) [9]. Let d_i and r_i be elements of 1-D domain and range blocks, respectively, we can obtain the minimal value rms_j occurred when s_j and o_j are optimized by means of

$$rms_j = \left[\sum_i \left(s_j \cdot d_i + o_j - r_i \right)^2 \right]^{\frac{1}{2}} \quad (4)$$

where i and j are pixel's order within a range block and the range block's order within a scanline, respectively. To determine the optimal values of s_j and o_j, the method of *least squares* is used. That is, the partial derivatives of square of Eq.(4) with respect to s_j and o_j are calculated, respectively, and then are set to zero so that the optimal values s_j and o_j are computed by using two linear equations. As the results, we obtain the optimal equations of parameters s_j and o_j as follows:

$$s_j = \frac{n\sum_i d_i r_i - \sum_i d_i \sum_i r_i}{n\sum_i d_i^2 - \left(\sum_i d_i \right)^2} \quad (5)$$

Material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and

and $$o_j = \frac{1}{n}\left(\sum_i r_i - s_j \sum_i d_i\right) \quad (6)$$

where n is a number of pixels in a range block. However, if the denominator of Eq.(5) is zero, then we define $s_j = 0$ and $o_j = \frac{1}{n}\sum r_i$. In practice, contractivity factor s_j of massic class is sensitive to convergence condition. Hence, to guarantee the contractivity criterion, s_j is limited by a maximum value, i.e., $|s_j| \le S_{max}$, where S_{max} is a maximum boundary of s_j. The details of convergence condition can be found in [11].

4.5 Bit allocation

Bit allocation is a process to appropriately define a number of bits for each encoded parameter so that the reconstructed image is close to the original one. In this paper, the following parameters, domain-block location (D_l), domain-block orientation (D_o), range-block size (R_s), scaling (s), and offset (o), represent the mutual fractal codes generated from a pair of contractive affine transformations and massic transformations. These parameters can be classified into two classes: geometric and massic classes according to their transformations [3], [12]. The geometric class is the outcomes from scanline partition and transformation and its bit allocation depends on geometrical characteristic of images, such as image size, range-block size, and others. Therefore, it is almost restricted. On the other hand, the massic class is the results of measuring similarity or distortion of domain-range mapping by means of root-mean-square metric. Thus, its bit allocation is very flexible and affects directly to the image quality. However, the fewer bits for encoded parameters (massic class) trade with the poorer image quality. For the suitable design, the bit requirement of mutual fractal codes is allocated as follows: For geometric class, domain-block location (D_l), domain-block orientation (D_o), and range-block size (R_s) are required 8, 1, 2 bits, respectively. For massic class, the scaling (s) and offset (o) parameters based on the optimal values are presented in [9]. These parameters require 5 and 7 bits, respectively. Hence, the total bit allocation for each range block can be obtained as follows:

$$B_{MSFC} = D_l + D_o + R_s + s + o \quad (7)$$

Note that in case of fixed-size partition, the factor R_s is disregarded, since all range-block sizes are equal. Therefore, the bit allocation for each range block is reduced by this factor.

5. Implementation and Application to Field-Sequential Stereoscopic Video

In this section, we have introduced the mutual-domain-pool design for both cases: one-dimensional (scanline) and two-dimensional implementation schemes of mutual fractal coding, and their application to field-sequential stereoscopic image sequences.

5.1 One-dimensional implementation scheme

The implementation scheme of mutual scanline fractal coding is illustrated in Fig. 8(a). We firstly describe the manipulation of the implementation scheme with the encoder section. In this scheme, the divide-and-conquer technique is used. The reason for applying this technique is given in Section 3. Thus, the dividing procedure is a preprocessing step that splits a whole image into a number of pairs of scanlines. Then, each pair of scanlines is divided by scanline partition. This scanline partition is an important procedure, since it defines the range-block size that is a significant factor to control bit rate and image quality. Therefore, in case of adaptive-size partition of the range blocks, the dynamic range threshold is a parameter used to control a number of range blocks. In other words, we can also control the bit rate of this scheme by using this threshold. The outcomes of this procedure are composed of two sets of nonoverlapping range blocks, i.e., R_L and R_R, and two sets of overlapping domain blocks, i.e., D_L and D_R. In this case, there are two domain pools (D_L- and D_R-domain pools) for keeping the sets of overlapping domain blocks and these domain pools are called *mutual domain pools*. After that all range blocks of R_L and R_R search for the matching blocks in the domain pool D_R and D_L by two domain-range-mapping algorithms $f : D_L \to R_R$ and $g : D_R \to R_L$, respectively. The domain-range-mapping algorithm consists of three components—geometric transformation, massic transformation, and bit allocation—for performing and generating mutual fractal codes. For clear explanation, the relationship of geometric and massic transformations is constructed. That is,

$$\begin{cases} f = (f_M \circ f_G)(x) = f_M(f_G(x)) \\ g = (g_M \circ g_G)(x) = g_M(g_G(x)) \end{cases} \quad (8)$$

where subscripts M and G denote the massic and geometric transformations, respectively. From Eq.(8), f_G and g_G are 1-D affine transformations as described in Subsection 4.1. To obtain the best matching of domain-range mapping, all domain blocks are shrunken their sizes to the range block sizes by pixel averaging. Then, each range block is matched with the shrunken domain block. In this step, the massic transformation is applied to measure the similarity of those domain and range blocks. In this case, the root-mean-square metric (Eq.(4)) is used and its minimum value is decided that which pair is the best matching of the domain-range mapping. Then, the scaling s and offset o are retained as a part of the mutual fractal codes. However, since s and o computed by Eqs.(5)-(6) are real and are infinite numbers, thus, they are quantized to be finite numbers. In this paper, the uniform quantization is applied and the suitable resolution is defined as given in Subsection 4.5. Finally, the merging procedure integrates the mutual fractal codes w_u and w_v from both domain-range-mapping algorithms into $W=\{D_l, D_o, R_s, s, o\}$, where these symbols denote domain-block location, domain-block orientation, range-block size, scaling, and offset, respectively.

For decoder section, the operation of each block inside the decoder can be described by Fig. 6(b). After decoding mutual fractal codes with a few iterations as depicted in Fig. 8(a), the reconstructed image appears.

5.2 Two-dimensional implementation scheme

The implementation scheme of mutual 2-D fractal coding (M2DFC) is depicted in Fig. 8(b). The main structure is

Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Copyright © by Chulalongkorn University
All rights reserved. Reproduction or transmission in whole or in part, any form or by any means, electronic, mechanical or otherwise, is prohibited without the prior written permission of the copyright owner.
Forbidden to modify the content, and cite the document when use.

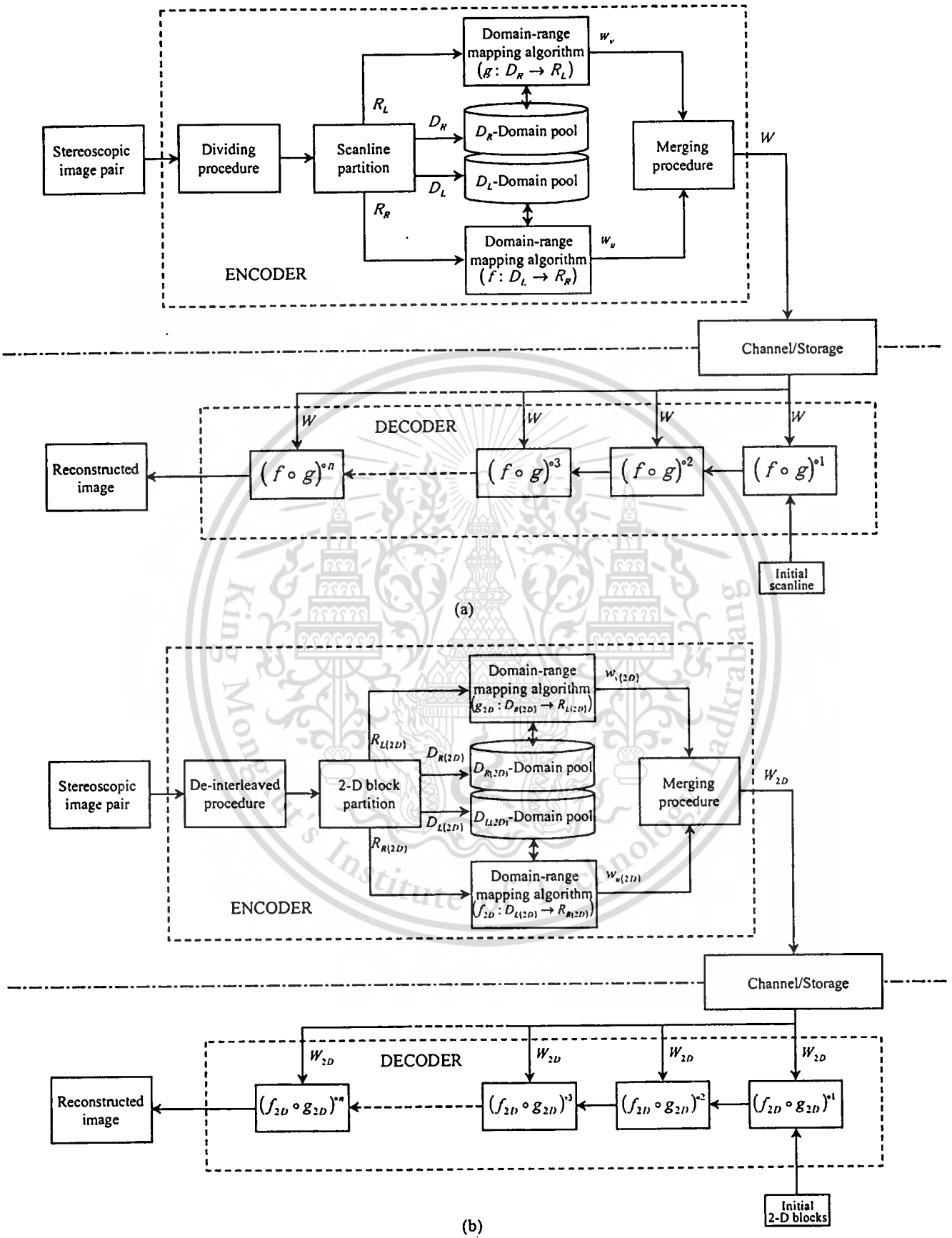


Fig. 8 Overall implementation schemes of (a) mutual scanline fractal coding and (b) mutual 2-D fractal coding

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and

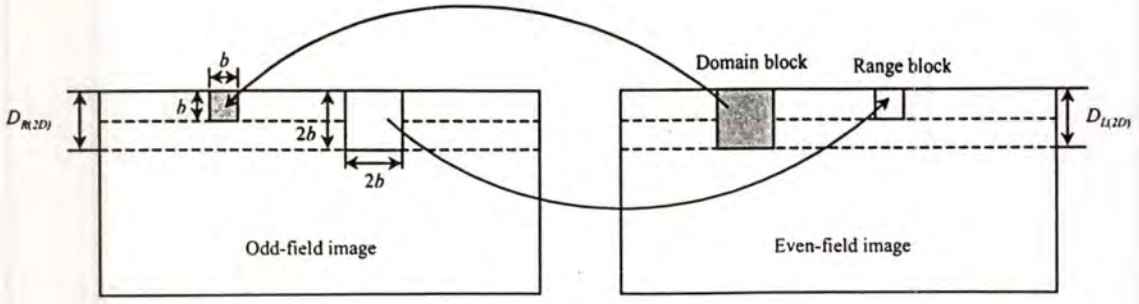


Fig. 9 The operation of mutual domain-range mapping

(a) (b)

Fig. 10 Original test images: (a) a "book" image and (b) a "chopstick" image

similar to the mutual scanline fractal coding (MSFC). All of designing phases in Section 4 can be extended easily to 2-D case. Thus, only the significant parts of the scheme are addressed.

In this scheme, the preprocessing step of encoder section is de-interleaved procedure. This procedure converts a frame -based image, containing both even and odd fields, into a pair of field-based images shown in Fig 3(b) so that the high vertical-frequencies of the image are eliminated. Then, the field-based images are partitioned into 2-D domain and range blocks. The outcomes from the 2-D block partition are $R_{L(2D)}$, $R_{R(2D)}$, $D_{L(2D)}$, and $D_{R(2D)}$, where subscripts L and R represent the left and right images. The $D_{L(2D)}$- and $D_{R(2D)}$-domain pools are the part of field images and are independent as shown in Fig. 9. In this case, the operation of domain-range mapping algorithms performs the same as in the scanline case. The different parts are that Eqs. (3) and (8) are replaced with Eqs. (9) and (10), respectively. Such that, Eq. (9) is

$$\begin{cases} f_{G_{2D}}(x_l, y_l) = \begin{bmatrix} l_1 & l_2 \\ l_3 & l_4 \end{bmatrix} \begin{bmatrix} x_l \\ y_l \end{bmatrix} + \begin{bmatrix} l_5 \\ l_6 \end{bmatrix}, & \forall x_l, y_l \in X_L \\ g_{G_{2D}}(x_r, y_r) = \begin{bmatrix} r_1 & r_2 \\ r_3 & r_4 \end{bmatrix} \begin{bmatrix} x_r \\ y_r \end{bmatrix} + \begin{bmatrix} r_5 \\ r_6 \end{bmatrix}, & \forall x_r, y_r \in X_R \end{cases} \quad (9)$$

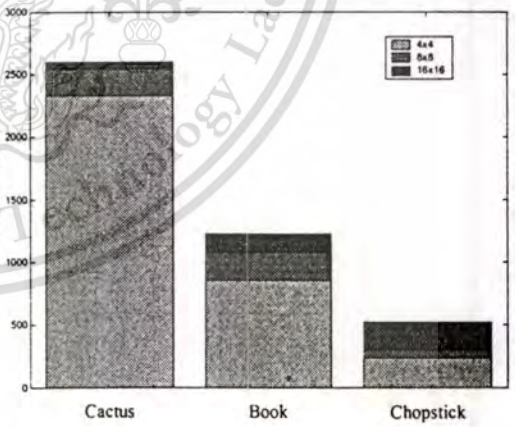


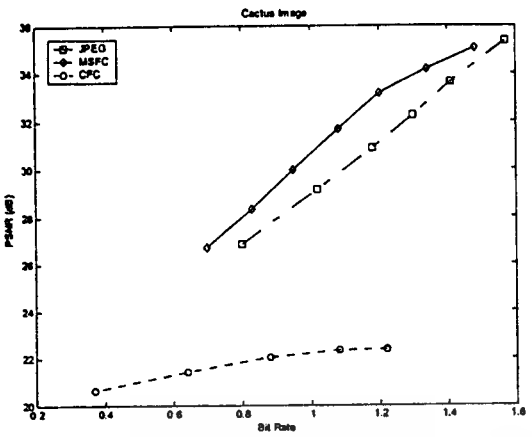
Fig. 11 Image-complexity comparison of the test images

where l_n and r_n denote the elements of geometrical contractive matrix, and $[x_l\ y_l]^T$ and $[x_r\ y_r]^T$ represent the pixel vectors of the left and right images, respectively, and Eq. (10) is

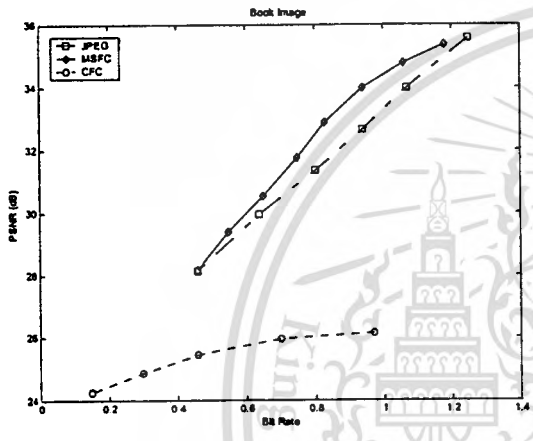
$$\begin{cases} f_{2D} = (f_{M_{2D}} \circ f_{G_{2D}})(x,y) = f_{M_{2D}}(f_{G_{2D}}(x,y)) \\ g_{2D} = (g_{M_{2D}} \circ g_{G_{2D}})(x,y) = g_{M_{2D}}(g_{G_{2D}}(x,y)) \end{cases} \quad (10)$$

where subscript $2D$ denotes 2-D affine transformations.

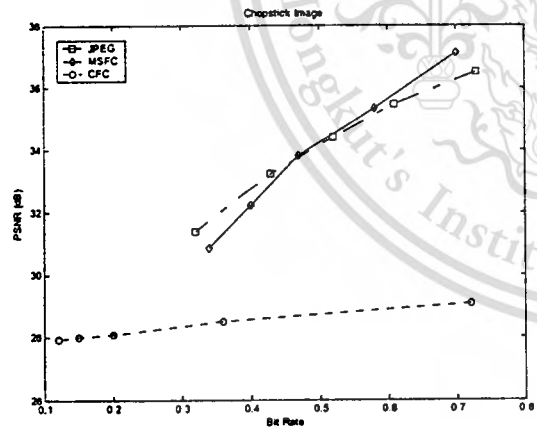
This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



(a)

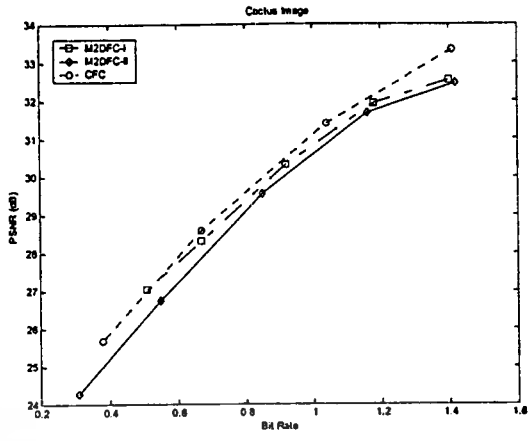


(b)

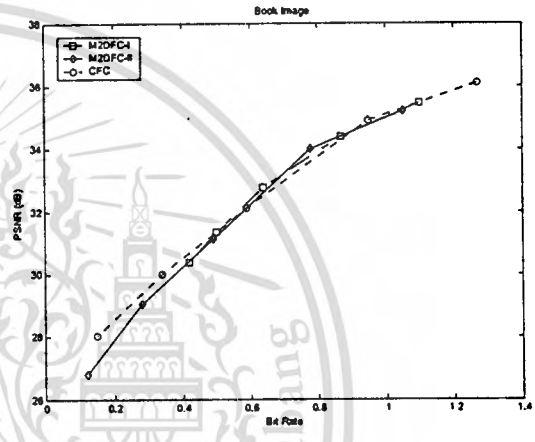


(c)

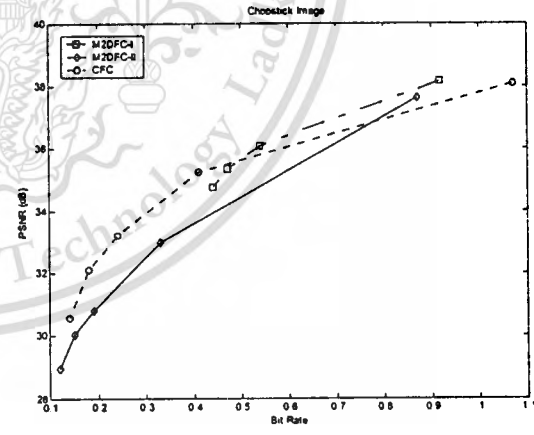
Fig. 12 Comparison of the frame-based reconstructed images of JPEG, CFC with10 iterations, and MSFC with 5 iterations



(a)



(b)



(c)

Fig. 13 Comparison of the field-based reconstructed images of CFC with 10 iterations, M2DFC-I, and M2DFC-II with 5 iterations

In addition, the bit requirement for mutual fractal codes of the M2DFC can be obtained as follows:

$$B_{M2DFC} = D_{l(2D)} + D_{o(2D)} + R_{s(2D)} + s_{2D} + o_{2D} \qquad (11)$$

where $D_{l(2D)}$, $D_{o(2D)}$, $R_{s(2D)}$, $s_{(2D)}$, and $o_{(2D)}$ denote domain-block location, domain-block orientation, range-block structure, scaling, and offset, respectively. In this case, it totally needs 27 bits (4 bits for range-block structure (three-level quadtree partition: 4x4, 8x8, and 16x16), 8 bits for domain-block location (column coordinate only), 3 bits for domain-block orientation, 7 bits for offset, and 5 bits for scaling, respectively) for each self-transformation, i.e. mutual fractal code. In this way, it is clear that this mutual-domain-pool design makes the coding system effective by the following reasons. (i) The domain-pool size is reduced, thus speeding

Forbidden to modify the content, and cite the document when use.

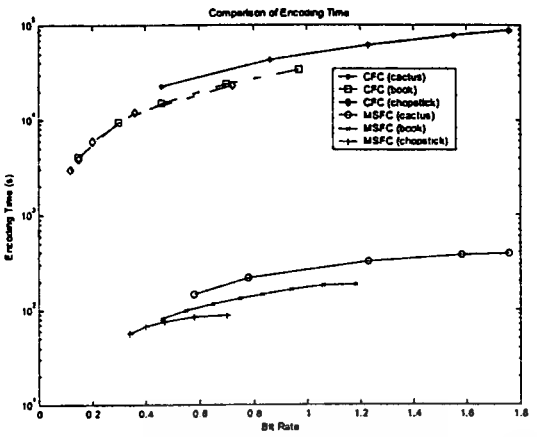


Fig. 14 Comparison of encoding time of CFC and MSFC in frame-based coding

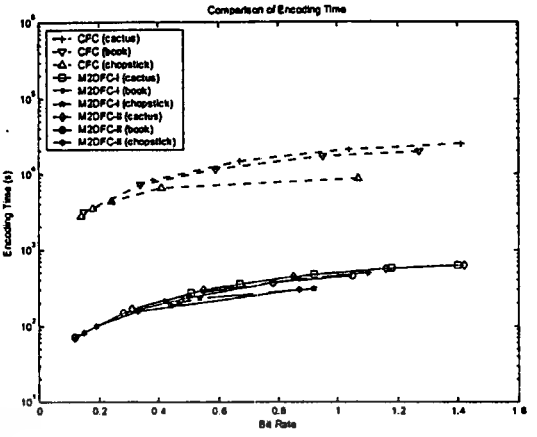


Fig. 15 Comparison of encoding time of CFC, M2DFC-I, and M2DFC-II in field-based coding

up the searching time while maintaining the image quality as good as possible. (ii) Saving the number of bits for storing the domain-block position, only the column coordinate is kept, whereas the row coordinate is omitted, since it implicitly is with the range block position. (iii) The mutual fractal codes generated from the mutual domain-range mapping algorithms provide the faster decoding of the reconstructed images for the reason in the Subsection 3.3.

6. Experimental Results

In this experiment, we testify and evaluate the performance of the intended codec in following issues: (i) encoding and decoding times and (ii) image quality. The condition of comparisons of the codec performance is based on fractal coding on spatial domain. The mutual scanline fractal coding (MSFC) and mutual 2-D fractal coding (M2DFC) algorithms are implemented and simulated by using MATLAB programming. The machine for performing the simulation program is the desktop PC with PENTIUM 4 processor and the clock speed is 2.4 GHz.

In order to testify and evaluate the performance of the intended and conventional codecs, we have prepared a set of test images, i.e., "cactus", "book", and "chopstick" images. The image quality is evaluated by means of peak signal-to-noise ratio (PSNR) defined as

$$\mathrm{PSNR} = 20\log\left(\frac{255}{\frac{1}{M\times N}(x-\hat{x})^2}\right) \qquad (12)$$

where x and $\hat{x}$ are the original and reconstructed images, respectively, and M and N are image dimensions. For coding efficiency, we evaluate it in the form of bit rate B_R that can be calculated by the following equation:

$$B_R = \frac{\text{Total size in bits of encoder output}}{\text{Total size in pixels of encoder input}} \qquad (13)$$

In case of the MSFC and M2DFC, the *total size in bits of encoder output* is the product of the total range blocks and their bit allocation. Hence, the bit allocation for each range block can be obtained by means of Eq.(7) for the MSFC and Eq.(11) for M2DFC.

Prior to experiment, the complexity of the test images is analyzed and summarized in Fig. 11. In this case, we use the variance (variant threshold = 0.3) of images to classify the complexity regions into three levels, 4x4-, 8x8-, and 16x16-blocks, respectively. These different square blocks, which imply that smaller blocks, mean more complicated regions. On the other hand, larger blocks mean more smooth regions. In Fig. 11, we can summarize that the "cactus" image is the highest complexity image, since it contains smaller square blocks (4x4 blocks) when compared to the other two images. The "book" and "chopstick" images are the median and low complexity images, respectively.

In the first experiment, we have evaluated the performance in the form of frame-based coding of the MSFC codec compared with the conventional fractal coding (CFC) and the JPEG standard coding codecs. In this experiment, the specifications of the MSFC and CFC codecs are set as follows. For the MSFC codec, the scanline technique is utilized to partition the range blocks with three-block-size adaptive partition (1x8, 1x16, and 1x32 block sizes). The first-match searching technique for domain-range mapping is selected. The bit requirement for each mutual fractal code is totally 23 bits as defined in Subsection 4.5. For the CFC codec, the basic and classical approach [9] with quadtree (three-level block sizes: 4x4, 8x8, and 16x16-blocks) partition is used. The image itself is assigned to be a domain pool. The searching technique is the same as the one that the MSFC is utilized. The bit allocation requires 33 bits for each fractal code. Then, the MSFC and CFC are applied to a set of test images, which are frame-based images. The fractal codes generated from both codecs are encoded by Huffman coding for the reduction of their redundancies and for the fairness of comparison with JPEG standard. The experimental results graphically illustrated in Fig. 12 indicate that in case of high and median complex images, i.e., "cactus" and "book" images, the MSFC codec is superior to both CFC and JPEG-standard codecs. For low complex image, the MSFC outperforms the JPEG for the bit rate greater than 0.45 bpp (See Fig. 12(c)). This is because discrete cosine transform

(DCT) carries out the better energy compaction for low complex images. In this case, for very low bit rate, only the DC coefficient is kept, thus reducing more bit rate. Simultaneously, this easily leads to the evident image crosstalk. However, in case of en-coding and decoding time, the MSFC is greatly improved when compared with the CFC codec as shown in Fig. 14 (see also Table 1).

Moreover, in order to evaluate the effectiveness of mutual domain pool design in 2-D case, the M2DFC codec is proposed and compared with the CFC codec. Prior to experiment, the following conditions are set for the fairness of comparison. In case of the CFC codec, the specifications are not changed from the first experiment. In case of the M2DFC, there are two versions of range block partition, M2DFC-I for two-level quadtree partition (4x4 and 8x8 range-block sizes) and M2DFC-II for three-level quadtree partition (4x4, 8x8, and 16x16 range-block sizes), respectively. For M2DFC-I codec, the bit requirement is 26 bits (3 bits for $R_{s(2D)}$), defined in subsection 5.2, for each mutual fractal code. For M2DFC-II codec, the bit allocation is 27 bits (4 bits for $R_{s(2D)}$). Then, the overall implementation scheme depicted in Fig. 8 (b) is applied to a set of test images, which are field-based images. The experimental results show that the quality of reconstructed images of the M2DFC-I, M2DFC-II, and CFC codecs is comparable as graphically exhibited in Fig. 13. For high complex image, the average degradation of reconstructed image quality of the M2DFC-I is lower than 0.15 dB at 0.92 bpp when compared with the CFC codec as illustrated in Fig. 13(a). For median complex image, it yields almost the same reconstructed images as the CFC codec. For low complex image, it outperforms the CFC codec for the bit rate greater than 0.47 bpp. In case of M2DFC-II, the average degradation is lower than 0.35 dB at 0.85 bpp for high complex image and 0.15 dB at 0.49 bpp for median complex image, and 1.30 dB at 0.19 bpp for low complex image, respectively, when compared with the CFC codec. Nevertheless, the encoding and decoding time of the M2DFC-I and M2DFC-II is greatly improved as shown in Fig. 15. Approximately, the mutual fractal coding is faster than the conventional fractal coding as summarized in Tables 2 and 3. This is a trade-off between the acceptable degrada-tion of reconstructed images and the significant improve-ment of mutual fractal coding.

Table 1 Comparisons of encoding time of MSFC and CFC

Test image	MSFC (s)	CFC (s)	Improved Ratio
Cactus	290.50	58,523.76	201
Book	139.82	17,388.55	124
Chopstick	74.94	9,552.05	127

Table 2 Comparisons of encoding time of M2DFC-I and CFC

Test image	M2DFC-I (s)	CFC (s)	Improved Ratio
Cactus	457.97	17,251.07	38
Book	337.54	11,676.23	35
Chopstick	231.01	5,137.30	22

Table 3 Comparisons of encoding time of M2DFC-II and CFC

Test image	M2DFC-II (s)	CFC (s)	Improved Ratio
Cactus	417.86	17,251.07	41
Book	254.22	11,676.23	46
Chopstick	139.79	5,137.30	37

7. Conclusions

In this paper, we have proposed the effective domain pool design, which is an important phase to speed up the encoding and decoding time of fractal coding. This design is called *mutual domain pool design* based on the equivalent property of two complete metric spaces. The mutual domain pool design leads to the effective coding system which can be summarized as follows: (i) The domain-pool size is reduced, thus decreasing the encoding time. (ii) Saving the number of bits for storing the domain-block position, only the column coordinate of domain-pool position is kept, whereas the row coordinate is omitted, since it implicitly is with the range-block position. (iii) The mutual fractal codes generated from the mutual domain-range mapping provide faster decoding of the reconstructed images. Furthermore, we have implemented the mutual domain pools for fractal coding in both 1-D and 2-D cases, i.e., MSFC and M2DFC schemes. To evidently evaluate the performance of the mutual domain pools without any bias, all schemes, including CFC codec, use the simple searching strategy, i.e., first-match searching technique, for domain-range mapping. The experimental results show that in case of frame-based images the MSFC is superior to the JPEG-standard and CFC codecs in terms of image quality as illustrated in Fig. 12. This is because the MSFC makes use of the scanline technique to divide the range blocks. Therefore, the high vertical-frequencies of the frame-based images are preserved. Furthermore, the MSFC greatly achieves the encoding and decoding time when compared with the CFC as demonstrated in Fig. 14. In case of field-based images, the quality of reconstructed images of M2DFC is acceptable degradation when compared with the CFC codec (see Fig. 13). In addition, the encoding and decoding time of the M2DFC is significantly improved as exhibited in Fig. 15 and Tables 2 and 3. We can summarize that the mutual domain pool design leads to a trade-off between the acceptable degradation of reconstructed image and the significant improvement of mutual fractal coding.

Although the mutual domain pool design can greatly improve the encoding and decoding time, some limitations still remained. For example, in case of a large image, the encoding and decoding time exponentially increases with the image size when using the simple searching technique, i.e., first-match searching technique. In this case, we are currently studying the extended version of the mutual domain pool design. That is the *multiple domain pool design*. Finally, we believe that the mutual domain pools and their extended concept are also able to collaborate with the other techniques—such as classifications [9], [13] feature vectors [13], frequency-like domains [14], [15], and others—for extreme-

This material is reserved for educational use. Forbidden to modify the content, and commercial use.

ly improving the fractal encoding and decoding time and image quality. This study will be reported in a future work.

Appendix

This appendix proves Theorems 1 and 2.

Proof of Theorem 1: By Definition 2, we have

$$d_R(f(x), f(y)) \le s_L \cdot d_L(x, y), \ \forall x, y \in X_L \quad \text{(A-1)}$$

and $$d_L(g(x), g(y)) \le s_R \cdot d_R(x, y), \ \forall x, y \in X_R \quad \text{(A-2)}$$

where s_L and s_R are contractivity factors for f and g, respectively. We begin with inequality (A-2) and multiply both sides by contractivity factor s_L. The result becomes

$$s_L \cdot d_L(g(x), g(y)) \le s_L \cdot s_R \cdot d_R(x, y) \quad \text{(A-3)}$$

However, we also have

$$d_R(f(g(x)), f(g(y))) \le s_L \cdot d_L(g(x), g(y)) \le s_L \cdot s_R \cdot d_R(x, y) \quad \text{(A-4)}$$

By the definition of contractive mapping, $s_L \cdot s_R$ is contractivity of $f \circ g$. This proof is now complete. □

Proof of Theorem 2: Let $X = X_L \cup X_R$ and $X_L \cap X_R = \varnothing$. Then, based on Definition 2, f and g are contractive transformations that mutually transform between two equivalent complete metric spaces (X_L, d_L) and (X_R, d_R). We want to prove that the sequence

$$(f \circ g)^{*n}(x) = \underbrace{f(g(f(g(\cdots(x)\cdots))))}_{n \text{ times}} \quad \text{(A-5)}$$

converges to a fixed point. By Theorem 1, let $s = s_L \cdot s_R$ and let $p > q$. Then for $x \in X$,

$$\begin{aligned} d\left((f \circ g)^{*q}(x), (f \circ g)^{*p}(x)\right) &< s \cdot d\left((f \circ g)^{*(q-1)}(x), (f \circ g)^{*(p-1)}(x)\right) \\ &< s^q \cdot d\left(x, (f \circ g)^{*(p-q)}(x)\right) \end{aligned} \quad \text{(A-6)}$$

Using triangle inequality, we obtain

$$\begin{aligned} d\left(x, (f \circ g)^{*(p-q)}(x)\right) &\le d\left(x, (f \circ g)^{*(p-q-1)}(x)\right) + \\ &\quad d\left((f \circ g)^{*(p-q-1)}(x), (f \circ g)^{*(p-q)}(x)\right) \\ &\le d(x, (f \circ g)(x)) + d\left((f \circ g)(x), (f \circ g)^{*2}(x)\right) \\ &\quad + \cdots + d\left((f \circ g)^{*(p-q-1)}(x), (f \circ g)^{*(p-q)}(x)\right) \\ &\le \left(1 + s + \cdots + s^{p-q-1}\right) d(x, (f \circ g)(x)) \\ &\le \frac{1}{1-s} d(x, (f \circ g)(x)) \end{aligned} \quad \text{(A-7)}$$

Substituting (A-7) to (A-6), we obtain

$$d\left((f \circ g)^{*q}(x), (f \circ g)^{*p}(x)\right) < \frac{s^q}{1-s} d(x, (f \circ g)(x)) \quad \text{(A-8)}$$

From Eq.(A-8), if p and q are sufficiently large and $s<1$, the left hand side of Eq.(A-8) is made as small as possible. This implies that the sequence

$$\{x, f(g(x)), f(g(f(g(x)))), \ldots\}$$

is a Cauchy sequence. Thus, the limit point

$$x_f = \lim_{n \to \infty} (f \circ g)^{*n}(x) \quad \text{(A-9)}$$

is in X, such that X is complete. Since $f \circ g$ is contractive as described in Theorem 1, it is continuous, and hence

$$(f \circ g)(x_f) = (f \circ g)\left(\lim_{n \to \infty} (f \circ g)^{*n}(x)\right) = \lim_{n \to \infty} (f \circ g)^{*(n+1)}(x) = x_f \quad \text{(A-10)}$$

Moreover, let x_{f1} and x_{f2} be two fixed points of $f \circ g$. Then

$$x_{f1} = (f \circ g)(x_{f1}), x_{f2} = (f \circ g)(x_{f2})$$

and

$$\begin{aligned} d(x_{f1}, x_{f2}) &= d\left((f \circ g)(x_{f1}), (f \circ g)(x_{f2})\right) \\ &\le s d(x_{f1}, x_{f2}) \end{aligned}$$

$$(1-s) d(x_{f1}, x_{f2}) \le 0 \quad \text{(A-11)}$$

which implies $d(x_{f1}, x_{f2}) = 0$, hence $x_{f1} = x_{f2}$. The proof is now complete. □

References

[1] T. Naemura, M. Kaneko and H. Harashima: Compression and representation of 3-D images, IEICE Trans. Inf. & Syst., Vol. E82-D, No. 3, pp. 558-567, 1999.

[2] M. Okui, A. Hanazato, F. Okano and I. Yuyama: A study on scanning methods for a field-sequential stereoscopic display, IEEE Trans. Circuits Syst. Video Technol., Vol. 10, No. 2, pp. 244-260, 2000.

[3] A. E. Jacquin: Image coding based on a fractal theory of iterated contractive image transformations, IEEE Trans. Image Processing, Vol. 1, No. 1, pp. 18-30, 1992.

[4] H. T. Chang and C. J. Kuo: Iteration-free fractal image coding based on efficient domain pool design, IEEE Trans. Image Processing, Vol. 9, No. 4, pp. 329-339, 2000.

[5] B. Wohlberg and G. de Jager: A review of the fractal image coding literature, IEEE Trans. Image Processing, Vol. 8, No. 12, pp. 1716-1729, 1999.

[6] A. J. Woods, D. Offszanka and G. Martin: A PC-based stereoscopic video walkthrough, Stereoscopic Displays and Virtual Reality Systems VI, Proceedings of The SPIE, pp. 306-312, San Jose, California, 1999.

[7] P. Limmaneepraserth and R. Varakulsiripunth: Encoding 3D video by using APOM, Proceedings of APCC/OECC, Vol. 2 of 2, pp. 849-852, Beijing, China, 1999.

[8] M. F. Barnsley: Fractals Everywhere Second Edition, Academic Press Professional, London, 1993.

[9] Y. Fisher: Fractal Image Compression, Theory and Application, Springer-Verlag, New York, 1995.

[10] J. Kominek: Advances in fractal compression for multimedia applications, Multimedia Systems, Vol. 5, No. 4, pp. 255-270, 1997.

[11] J. Kominek: Convergence of fractal encoded images, Proceedings of DDC, pp. 242-251, Snowbird, UT, USA, 1995.

[12] A. E. Jacquin: Fractal image coding: a review, Proc. IEEE, Vol. 81 No. 10, pp. 1451-1465, 1993.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

[13] M. Polvere and M. Nappi: Speed-up in fractal image coding: comparison of methods, IEEE Trans. Image Processing, Vol. 9, No. 6, pp. 1002-1009, 2000.
[14] T. -K. Truong, J. -H. Jeng, I. S. Reed, P.C. Lee and A. Q. Li: A fast encoding algorithm for fractal image compression using the DCT inner product, IEEE Trans. Image Processing, Vol. 9, No. 4, pp. 529-535, 2000.
[15] J. H. Jeng, T. K. Truong and J. R. Sheu: Fast fractal image compression using the Hadamard transform, IEE Proc.-Vis. Image Signal Process., Vol. 147, No. 6, pp. 571-574, 2000.

Ruttikorn Varakulsiripunth received his B.E. degree in electrical and electronics from Kyoto University, Kyoto, Japan in 1978. He obtained his M.E. and Ph.D. degrees in electrical and communication engineering from Tohoku University, Sendai, Japan in 1983 and 1986, respectively. He is now an associate professor in the Department of Electronics, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand. His current research interests are concerned with computer communication network including switching system, queueing analysis, flow and congestion control, multimedia communication, wireless communication, image processing and natural language processing.

Para Woraratpanya received his B.Ind. Tech in computer technology and M.Eng in computer engineering in 1992 and 1996, respectively, both from King Mongkut's Institute of Technology Ladkrabang. He is a lecturer in the Department of Computer Education, Faculty of Technical Education, King Mongkut's Institute of Technology North Bangkok. He is currently working for Ph.D degree in electrical engineering at King Mongkut's Institute of Technology Ladkrabang. His research interests include stereoscopic acquisition and compression, multimedia coding and processing, and digital signal processing.

(Received February 23, 2004; revised August 23, 2004)

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and

Author Biography

PERSONAL INFORMATION

Last Name:	WORARATPANYA (LIMMANEEPRASERTH)
First Name:	Kuntpong (Para)
Date of Birth:	02 October 1966
Nationality:	Thai
Office Address:	Department of Computer Education, Faculty of Technical Education, KMITNB 1518 Piboolsongkram Rd., Bangsue, Bankok 10800

EDUCATION & QUALIFICATION

1990-1992:	B.Ind.Tech. in Computer Technology (Second Class Honor), Faculty of Engineering, KMITL
1993-1996:	M. Eng in Electrical Engineering, Faculty of Engineering, KMITL

CURRENT EMPLOYMENT

1998-Present:	Lecturer, Department of Computer Education, KMITNB
Course Responsibilities:	1. Data Structures and File Processing 2. Digital Signal Processing 3. Multimedia Coding and Processing 4. Information Theory

RESEARCH AREAS

1. Multimedia Coding
2. Video Coding
3. Data Compression
4. Stereoscopic Acquisition and Compression

INTERNATIONAL PUBLICATIONS (1999-PRESENT)

[1] **Para Limmaneepraserth** and Ruttikorn Varakulsiripunth, "Encoding 3D Video by Using APOM," *Proceedings of APCC/OECC* (*Fifth Asia-Pacific Conference on Communications and Fourth Optoelectronics and Communications Conference*), Vol. 2 of 2, pp. 849-852, Beijing, China, 18-22 Oct. 1999.

[2] **Para Limmaneepraserth** and Ruttikorn Varakulsiripunth, "Generating the Stereoscopic Video by Using Active/Passive Optical Mechanism," *Proceedings of ISPACS'99* (*1999 IEEE International Symposium on Intelligent Signal Processing and Communication Systems*), pp 343-346. Phuket, Thailand, 8-10 Dec. 1999.

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

INTERNATIONAL PUBLICATIONS (1999-PRESENT)

[3] Chompoonoot Tengcharoen, Jirapat Kittimethee, **Para Limmaneepraserth**, and Ruttikorn Varakulsiripunth, "Image Ghosting Elimination in an I-frame of Interlaced 3D Video," *Proceedings of ICOM'01 (The First International Conference on Mechatronics)*, Vol. 2 of 2, pp. 409-418, Kuala Lumpur, Malaysia, 12-13 Feb. 2001.

[4] Patiroop Yuanghirun, Jeeragone Inrut, Sarayut Paludkong, Supot Nitsuwat, and **Para Limmaneepraserth**, "Thai Word Segmentation Using Combination of Forward and Backward Longest Matching Techniques," *Proceedings of ISCIT 2001 (International Symposium on Communications and Information Technology)*, pp. 37-40, Chiang Mai, Thailand, 14-16 Nov. 2001.

[5] Jirapat Kittimethee, **Para Limmaneepraserth**, and Ruttikorn Varakulsiripunth, "Error Concealment for Stereoscopic Video," *Proceedings of ROVISP 2003 (the International Conference on Robotics, Vision, Information and Signal Processing)*, pp. 39-44, Penang, Malaysia, 22-24 Jan. 2003.

[6] **Para Limmaneepraserth** and Ruttikorn Varakulsiripunth, "Designing and Implementing Mutual Scan-line Fractal Coding for Field-sequential Stereo Video Sequences," *Proceedings of ISSPA2003 (Seventh International Symposium on Signal Processing and Its Applications)*, pp. 109-112, Paris, France, 1-4 Jul. 2003.

[7] **Para Limmaneepraserth** and Ruttikorn Varakulsiripunth, "Study of Mutual Scan-line Fractal Coding," *Proceedings of VCIP2003 (Visual Communications and Image Processing 2003)*, pp. 2061-2070, Lugano, Switzerland, 8-11 Jul. 2003.

[8] Sanya Khruahong, Supot Nitsuwat, and **Para Limmaneepraserth**, "Thai Syllable Segmentation for Text-to-Speech Synthesis by Using Suited-syllable-structure Mapping," *Proceedings of ISCIT2003 (Third International Symposium on Communications and Information Technologies)*, pp. 948-952, Songkhla, Thailand, 3-5 Sept. 2003.

[9] Ruttikorn Varakulsiripunth and **Para Woraratpanya**, "Mutual Fractal Coding for Field-Sequential Stereoscopic Video," *Journal of Signal Processing*, pp. 165-178. Vol. 9 No. 2, Mar. 2005.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.