ECHO SIGNALS ANALYSIS OF MONOSTATIC RADAR

SOMKIAT LERKVARANYU

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF ENGINEERING IN ELECTRICAL ENGINEERING
SCHOOL OF GRADUATE STUDIES
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2003

ISBN 974-324-862-5

หัวข้อวิทยานิพนธ์	การวิเคราะห์สัญญาณสะท้อนกลับของระบบเรดาร์
นักศึกษา	นาย สมเกียรติ ฤกษ์วรัญญู
รหัสประจำตัว	39060062
ปริญญา	วิศวกรรมศาสตรดุษฎีบัณฑิต
สาขาวิชา	วิศวกรรมไฟฟ้า
พ. ศ.	2546
อาจารย์ผู้ควบคุมวิทยานิพนธ์	รศ.ดร.กอบชัย เดชหาญ
อาจารย์ผู้ควบคุมวิทยานิพนธ์ร่วม	รศ.ดร.ฟูศักดิ์ ชีวสุวิทย์

บทคัดย่อ

การกระเจิงและการลดทอนของแสง สามารถนำมาใช้ในการศึกษาสภาพชั้นบรรยากาศ โดยอาศัยเครื่องมือวัดระยะไกลของแสง ซึ่งจะถูกเรียกว่าเลเซอร์เรดาร์ที่ใช้ในการศึกษาสภาพชั้นบรรยากาศ โดยระบบเลเซอร์เรดาร์จะทำการส่งแสงเลเซอร์ไปยังชั้นบรรยากาศเพื่อทำการวัดอนุภาคต่างๆ เช่น ไอน้ำ ผลึกน้ำแข็ง สารผสมเอรอซอล และ ก๊าซ ชนิดต่างๆ การศึกษาและทดลองของเลเซอร์เรดาร์ที่ติดตั้งอยู่บนพื้นดินเพื่อทำการวัดหาบริเวณที่ปรากฏสารผสมเอรอซอล โดยแสดงอยู่ในรูปอัตราส่วนในการสะท้อนระหว่างการกระเจิงของแสงแบบ Mie และการกระเจิงของแสงแบบ Rayleigh ได้ถูกนำเสนอในวิทยานิพนธ์นี้ โดยอาศัยทฤษฎีการกระเจิงของ Rayleigh และรูปแบบความหนาแน่นในชั้นบรรยากาศ สามารถที่จะหาพื้นที่การกระเจิงของแสงที่เกิดจากสารผสมเอรอซอลได้ การแสดงบริเวณที่มีสารผสมเอรอซอลจะแสดงอยู่ในรูปอัตราส่วน ของการสะท้อนของสัญญาณ ที่วัดได้กับสัญญาณที่เกิดจากอนุภาคของก๊าซในชั้นบรรยากาศเท่านั้น และภายหลังจากการวัดและประมวลผลข้อมูล ระบบการจัดข้อมูลโดยอาศัยเทคนิคการเรียนรู้ข้อมูลด้วยตนเองได้ถูกนำเสนอในวิทยานิพนธ์นี้ด้วย

Thesis Title	Echo Signals Analysis of Monostatic Radar
Student	Mr. Somkiat Lerkvaranyu
Student ID.	39060062
Degree	Doctor of Engineering
Programme	Electrical Engineering
Year	2003
Thesis Advisor	Assoc.Prof.Dr. Kobchai Dejhan
Thesis Co-Advisor	Assoc.Prof.Dr. Fusak Cheevasuvit

ABSTRACT

Light scattering and attenuation can be used to investigate the atmosphere using an optical remote sensing instrument, which is called the atmospheric laser radar. Laser radar uses laser pulses to measure atmospheric constituents such as aerosol particles, ice crystals, water vapor, or trace gases. The study and experiments of ground-based laser radar to locate the aerosol distributions based on scattering ratio between Mie backscattering and Rayleigh backscattering are presented in this thesis. Rayleigh scattering theory and an atmospheric density model, the measured molecular signal provides a known target at each altitude from which the aerosol backscatter cross section can be determined. The aerosol profiles in the form of scattering ratio can also be determined by comparing the measured signal to the molecular signal calculated in the absence of aerosol extinction. Moreover, for the post processing, the atmospheric laser radar data classification system based on an unsupervised technique is also provided in this thesis.

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Kobchai Dejhan for providing me the opportunity to work with some of the most unique and existing lidar instruments. I also appreciate all the time he has spent guiding me through this part of my education.

My thanks also go to Professor Dr. Yoshikazu Miyanaga for his thoughtful and constructive comments about the unsupervised neural network classification system in this thesis to help improve the evaluation of this work.

I would also like to thank all the members of the optical remote sensing section of CRL Lidar Group who have helped make this work possible.

Finally, I would like to thank my family for their patience and support during this part of my education.

Somkiat Lerkvaranyu

CONTENTS

CONTENTS

FIGURE CONTENTS

FIGURE CONTENTS

TABLE CONTENTS

Chapter 1
Introduction

1.1 Preliminaries

To study the behavior of the atmosphere, the various laser instruments have been applied after the laser was invented. The atmosphere is very complicated system. The laser has been used to study the composition of mixing particles in the atmosphere is called atmospheric laser radar or lidar (Light Detection and Ranging). Currently, there are many types of lidar systems which are used to study the mixing complex particles in the atmosphere, such as the Rayleigh lidar, Mie lidar, Differential absorption lidar, Raman lidar and so on. All systems of lidar use the various laser properties to interact with particle concentrations in the atmosphere based on the absorption, scattering, depolarization, or Doppler frequency shift. Moreover, Mie and Rayleigh lidar techniques will be concern with elastic scattering between light and molecules. In this thesis, the ground-based lidar has been installed at King Mongkut's Institute of Technology Ladkrabang under collaboration with Communication Research Laboratory (CRL) of JAPAN. The purpose of this system is used to study and measure the aerosol distributions over 3 km in the boundary layer. The common technique of a aerosol profile is the scattering ratio involving both Rayleigh scattering from molecules and Mie scattering from other particle or aerosols in the atmosphere. The measurements were carried out in the nighttime

1.2 Thesis outlines

Chapter 2 provides the various basic backgrounds such as radar system, radar equation, and the characteristic of the atmosphere. The gases in the atmosphere interact with solar irradiation and with radiation from the earth's surface. Also in this chapter will mainly introduce the fact that the atmosphere can have a profound effect (reflection, scattering,

absorption and transmission) on intensity and spectral composition of the radiation. The various types of remote sensors and lidar are also provided in this chapter.

The applications with experimental results of laser radar or lidar for measuring the aerosol distributions, which will be discussed in the chapter 3. In the experiments, the particles in the atmosphere are measured by using the second harmonics of Nd : YAG laser (532 nm). The laser source is installed in the vertical axis with respect to the ground surface. A telescopes is used to receive the scattered laser wave in the sky.

In chapter 4, the matters focus on the cluster validity that is used in the process of data classification. A cluster validity criterion is the Total Fitting Level method (TFL) will be discussed. The concept of TFL is based on the difference between an actual data histogram and the estimated cluster probability distribution. The TFL will be used with the clustering algorithm for classification of the atmospheric laser radar data in the next chapter.

Chapter 5 provides a clustering and indexing system for atmospheric laser radar data based on an unsupervised neural network. This system adopts the self-organizing clustering (SOC). During learning, the total fitting level (TFL) method is used for cluster validation. The validity index function TFL(c) gives the minimum value at the optimal cluster number. By means of the TFL method, the system can learn by itself to find not only the optimal cluster number but also the exact value of two unknown parameters. Finally the system will be used for classification of aerosols distributions. Conclusions are given in chapter 6.

Chapter 2

Radar System

2.1 Introduction

Radar is an acronym for Radio Detection and Ranging and operates as an active sensor system. It almost operates in part of the microwave region of the electromagnetic spectrum, specifically in the frequency interval from 40 GHz to 300 MHz. The latter frequency extends into the higher frequencies of the broadcast-radio region. Unlike passive sensors that sense radiation from targets illuminated by the Sun or thermal sources, radar generates its own illumination by sending bursts or pulses of electromagnetic energy that reflect off of a target. A fraction of the reflected energy then returns to the radar's receiving antenna, which collects it and passes it to the electronic system. Thus, a radar system is a ranging device that measures distances as a function of round trip travel times in the atmosphere with speed of light. We can also derive information about target shapes and certain diagnostic physical properties of materials by analyzing signal modifications.

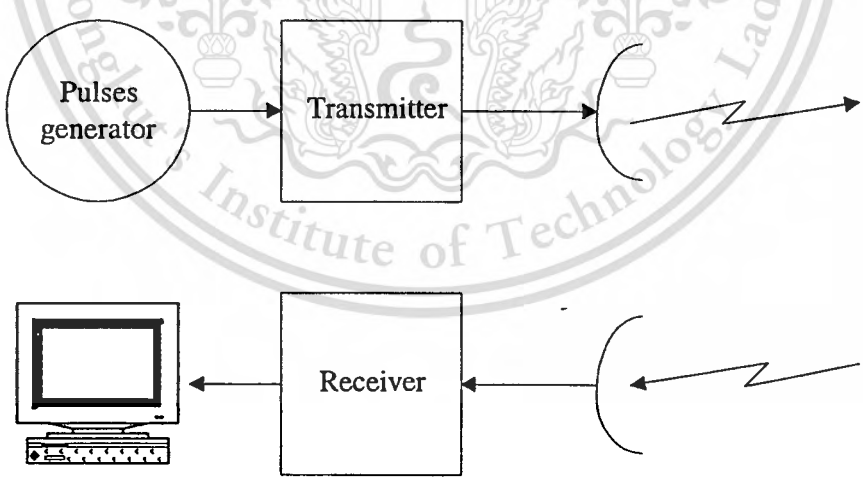


Figure 2.1 Radar system

In Fig. 2.1, a typical radar system consists of the following components:

(1) Pulse generator that generates timed pulses of microwave/radio energy

(2) Transmitter

(3) Directional antenna that shapes and focuses each pulse into a stream

(4) Receive antenna picks up the echo signals and sends to a receiver.

(5) Recording device which stores them digitally for later processing, and real time analog display on a cathode ray tube (CRT).

2.2 Radar Equation

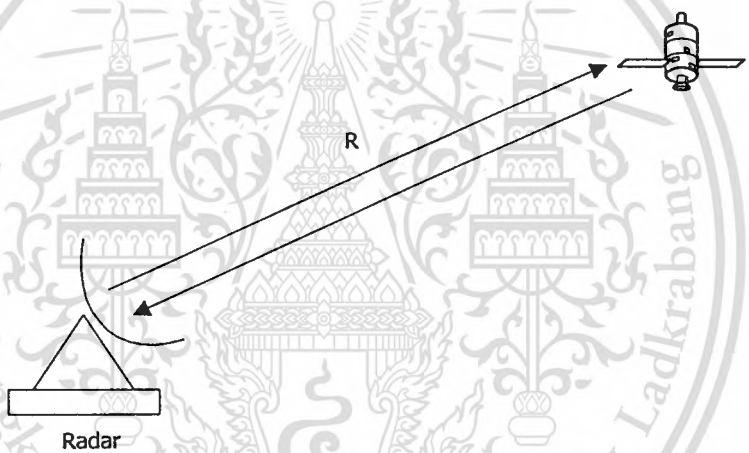


Figure 2.2 Radar operation

When the power radiated an isotropic denoted as P. In Fig. 2.2, at a distance R the power density P_d is

$$P_d = \frac{P}{4\pi R^2} \tag{2.1}$$

When the antenna gain G_t is incorporated into Eq. 2.1 , the power density in a particular antenna pointing direction P_t is

$$P_i = \frac{PG_t}{4\pi R^2} \tag{2.2}$$

The reflected power P_{re} with the target cross section σ becomes

$$P_{re} = \frac{PG_t}{4\pi R^2}\sigma \tag{2.3}$$

Power received at antenna P_r will be obtained

$$P_r = \frac{PG_t}{4\pi R^2}\sigma\frac{A_e}{4\pi R^2} \tag{2.4}$$

A_e is the effective area of the antenna which relates to antenna gain as

$$A_e = \frac{G_r\lambda^2}{4\pi} \tag{2.5}$$

If the same antenna is used for both transmitting and receiving, then G_t=G_r. The received power P_r is obtained as

$$P_r = \frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 R^4} \tag{2.6}$$

All parameters in this formula except "σ" is determined by the system. Only σ is a parameter related to the target.

When the energy reaches the target, some of the energy is reflected back towards the sensor. This backscattered microwave radiation is detected, measured, and timed. The time period can be measured from a transmitted signal travelling through the air reaching the target and being scattered back to the antenna. We then can determine the distance between the antenna and the target.

$$R = \frac{ct}{2} \tag{2.7}$$

where R is the target range, c is the speed of light and t is the round-trip time period for a transmitted pulse.

By supplying its own illumination, radar can function day and night without significant interference from blocking atmospheric conditions such as clouds, rain, etc., for some wavelengths. These characteristics prompted radar development in World War II for tracking aircraft and ships. Ground and airborne radar systems are used extensively today for marine navigation and air traffic control. Imaging radar, mounted on air or space platforms, has proven especially useful in mapping cloud-shrouded land surfaces. This use also permits expressing surface shapes in regions heavily covered by vegetation. In this thesis the radar is considered as a remote sensing system that is specially used for studying the behavior of the atmosphere.

2.3 Atmosphere

The atmosphere is a mixture of gases with constant proportions up to 80 km or more from ground. The main atmospheric gases are 78 percent nitrogen, 21 percent oxygen, and 1 percent argon. The atmosphere also contains water vapor, with concentrations varying with latitude and season, and a number of other trace gases, including carbon dioxide, methane, carbon monoxide, oxides of nitrogen, and ozone. Carbon dioxide is the principal atmospheric gas, with its concentration varying with time. It is increasing since the beginning of this century due to the burning of fossil fuels. Air is highly compressible. Half of its mass occurs in the lowest 5 km. Scientists defined five atmospheric regions or layers based on whether the temperature is increasing or decreasing, which is shown in Fig. 2.3, within the layer. Note that the temperatures and conditions in the atmosphere vary over the course of years, months, and even days. So, the extent of the layers varies with time. The regions or layers of the atmosphere include, from lowest to highest, the Troposphere, the Stratosphere, the Mesosphere, the Thermosphere, and the Exosphere.

The height of the layers varies from the equator to the poles and from day to day.

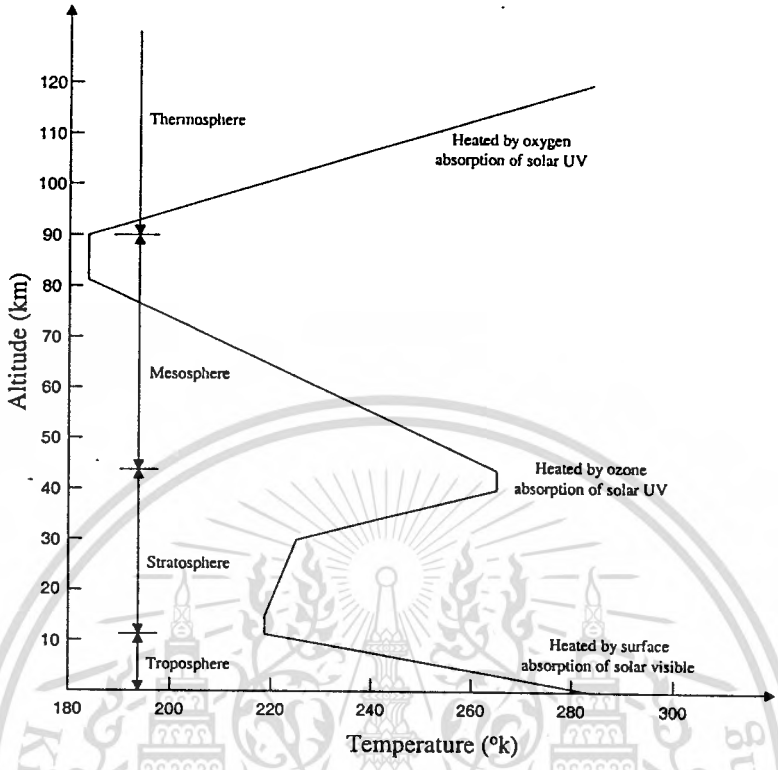


Figure 2.3 Characteristic of temperature in the atmosphere

2.4 Interactions with the Atmosphere

Most remote sensing is conducted above the earth either within or above the atmosphere. The gases in the atmosphere interact with solar irradiation and with radiation from the earth's surface. Although the incoming irradiation is a single source of excitation of atoms and molecules in the air and any materials found at the surface, that electromagnetic irradiation would experience varying degrees of transmission, absorption, and/or scattering depending on whatever wavelengths are considered. Figure 2.4 shows the processes of atmospheric radiation.

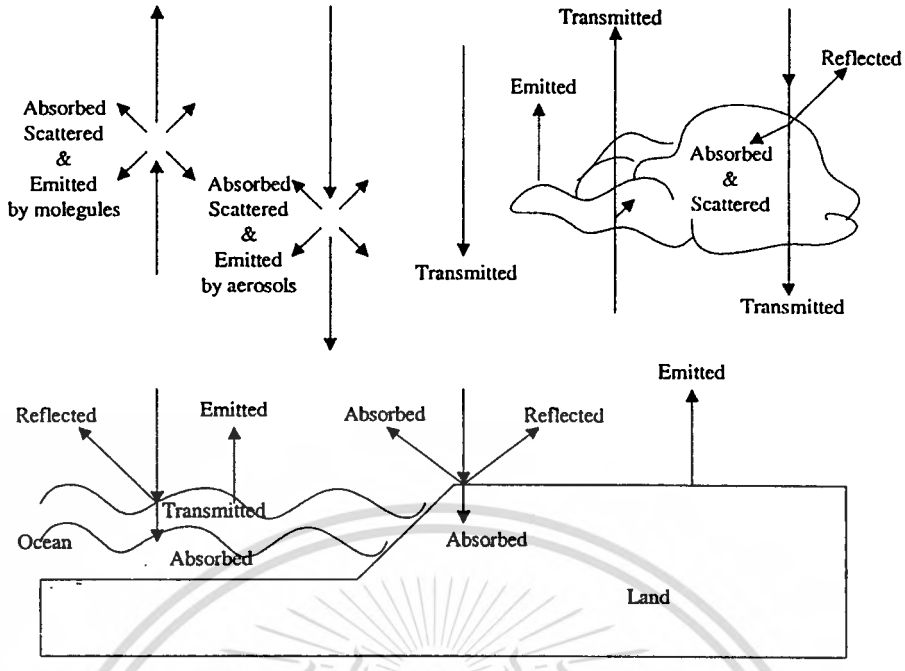


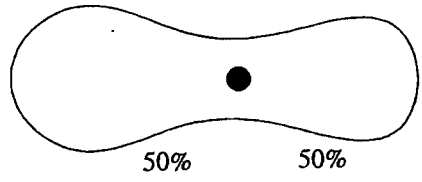
Figure 2.4 Processes of atmospheric radiation

The atmosphere has different effects on the electromagnetic transfer at different wavelengths. In this section, we will mainly introduce the fact that the atmosphere can have a profound effect on intensity and spectral composition of the radiation that reaches a remote sensing system. These effects are caused primarily by the atmospheric scattering and absorption.

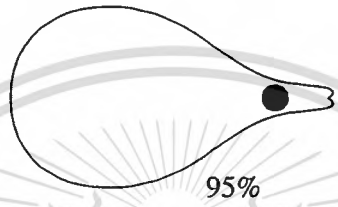
2.4.1 Scattering

Scattering occurs when particles or large gas molecules present in the atmosphere interact with and cause the electromagnetic radiation to be redirected from its original path. How much scattering takes place depends on several factors including the wavelength of the radiation, the abundance of particles or gases, and the distance the radiation travels through the atmosphere. There are three types of scattering which take place. The redirection of electromagnetic energy by the suspended

particles in the air with different particle sizes will have different effects on the electromagnetic energy propagation.



Rayleigh scattering



Mie scattering

Figure 2.5 Rayleigh and Mie scattering

In Fig. 2.5, Rayleigh scattering occurs when particles are very small compared to the wavelength of the radiation, first type of scattering. These could be particles such as small specks of dust or nitrogen and oxygen molecules. Rayleigh scattering causes shorter wavelengths of energy to be scattered much more than longer wavelengths. Rayleigh scattering is the dominant scattering mechanism in the upper atmosphere. The fact that the sky appears blue color during the day is because of this phenomenon. As sunlight passes through the atmosphere, the shorter wavelengths (i.e. blue) of the visible spectrum are scattered more than the other longer of visible wavelengths. At sunrise and sunset the light has to travel farther through the atmosphere than at midday and the scattering of the shorter wavelengths is more complete; this leaves a greater proportion of the longer wavelengths to penetrate the atmosphere.

Mie scattering, which is the second type of scattering, occurs when the particles are just about the same size as the wavelength of the

radiation. Dust, aerosol, smoke and water vapor are common causes of Mie scattering which tends to affect longer wavelengths than those affected by Rayleigh scattering. Mie scattering occurs mostly in the lower portions of the atmosphere where larger particles are more abundant, and dominates when cloud conditions are overcast. The last scattering mechanism of importance is called nonselective scattering. This occurs when the particles are much larger than the wavelength of the radiation. Water droplets and large dust particles can cause this type of scattering. Nonselective scattering gets its name from the fact that all wavelengths are scattered about equally.

2.4.2 Absorption

Atmosphere selectively absorbs energy in different wavelengths with different intensity. The atmosphere is composed of N_2, O_2, CO_2, H_2O, CO, SO_2, etc. Since different chemical element has a different spectral property, regions with different intensity. As a result, the atmosphere has the combined absorption features of various atmospheric gases. Figure 2.6 shows the major absorption wavelengths by CO_2, H_2O, O_2, O_3 in the atmosphere. Here is a generalized diagram showing relative atmospheric radiation transmission and absorption at different wavelengths. Dark zones mark minimal passage of incoming and/or outgoing radiation, whereas, white areas denote atmospheric windows, in which the radiation does not interact much with air molecules and hence, is not absorbed.

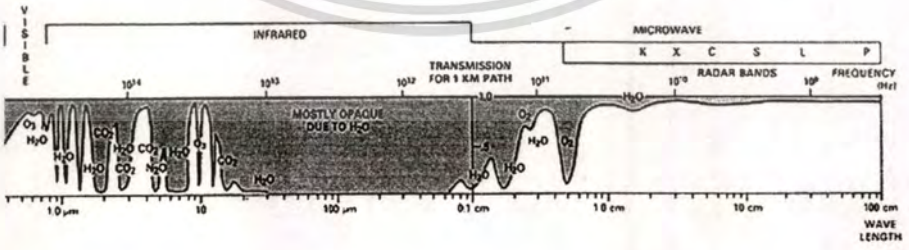


Figure 2.6 Absorption in the atmosphere

2.4.3 Transmission

The remaining amount of energy after being absorbed and scattered by the atmosphere is transmitted.

$$\text{Transmission} = \frac{\text{Transmitted Energy}}{\text{Input Energy}}$$

H_2O is most variable in the atmosphere. CO_2 varies seasonally. Therefore, the absorption and/or transmission of electromagnetic energy by H_2O and CO_2 are the most difficult part to be characterized.

2.4.4 Atmospheric Window

Atmospheric absorption reduces the number of spectral regions that some sensor systems can work with in observing the earth. It affects the decision in selecting and designing sensor. It has to consider

1) the spectral sensitivity of sensors available;
2) the presence and absence of atmospheric windows;
3) the source, magnitude, and spectral composition of the energy available in these ranges.

For the third point, there are to more consider the decision of choosing sensors and spectral regions on the manner in which the energy interacts with the target under investigation. On the other hand, although certain spectral regions may not be as transparent as others, they may be important spectral ranges in the remote sensing of the atmosphere. Figure 2.7 shows the atmospheric window.

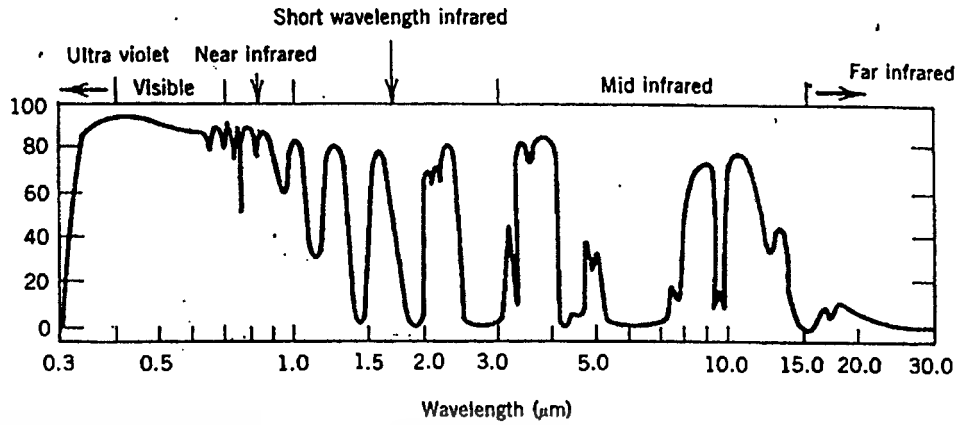
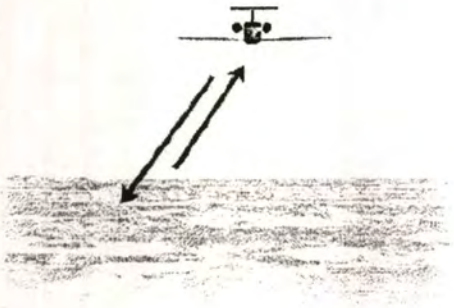


Figure 2.7 Atmospheric window

2.5 Remote Sensing

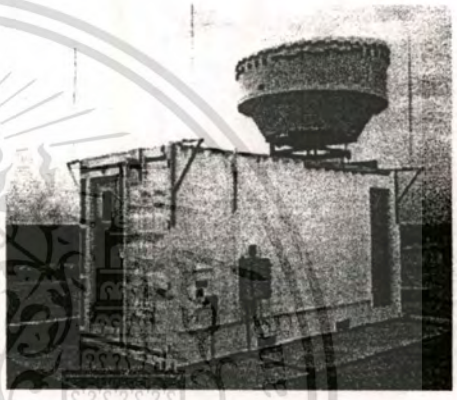
Remote sensing is a technology used for obtaining information about a target through the analysis of data acquired from the target at a distance. They are like cameras except that they use not only visible light but also other bands of the electromagnetic spectrum such as infrared, microwave and ultraviolet. Because they are so high up, these sensors can make images of a very large area. Basically, it is composed of three parts; the targets, the data acquisition which certain instruments, and the data analysis. This definition is so broad that the vision system of human eyes, sonar sounding of the sea floor, ultrasound and x-rays used in medical sciences and laser probing of atmospheric particles, are all included. The target can be as big as the earth, the moon and other planets. There are various remote sensing techniques for information extraction about the identity, quantity, spatial and temporal distribution of various targets of interest. These systems can be conducted on such platforms as aircraft, satellites, balloons, rockets, space shuttles, etc., which show in the Fig. 2.8. Various sensors are used to collect data include aerial photographic cameras and non-photographic instruments, such as radiometers, electro-optical scanners, radar systems, etc.



Airborne system

Satellite remote sensing

Spaceborne lidar



Ground-based rain radar

Figure 2.8 Remote sensing systems

2.6 Passive vs. Active Sensing

Figure 2.9 shows the various types of sensor. The two broadest classes of sensors are *Passive* and *Active.* In subdivision the sensors can be *non-imaging systems*, which measure the radiation received from all points in the sensed target and reports the result as an electrical signal, or *imaging* systems, which use a film or an image display device like a TV, computer monitor or a cathode ray tube. Since the radiation is related to specific points in the target, the end result is the image or the raster display in the parallel and horizontal lines.

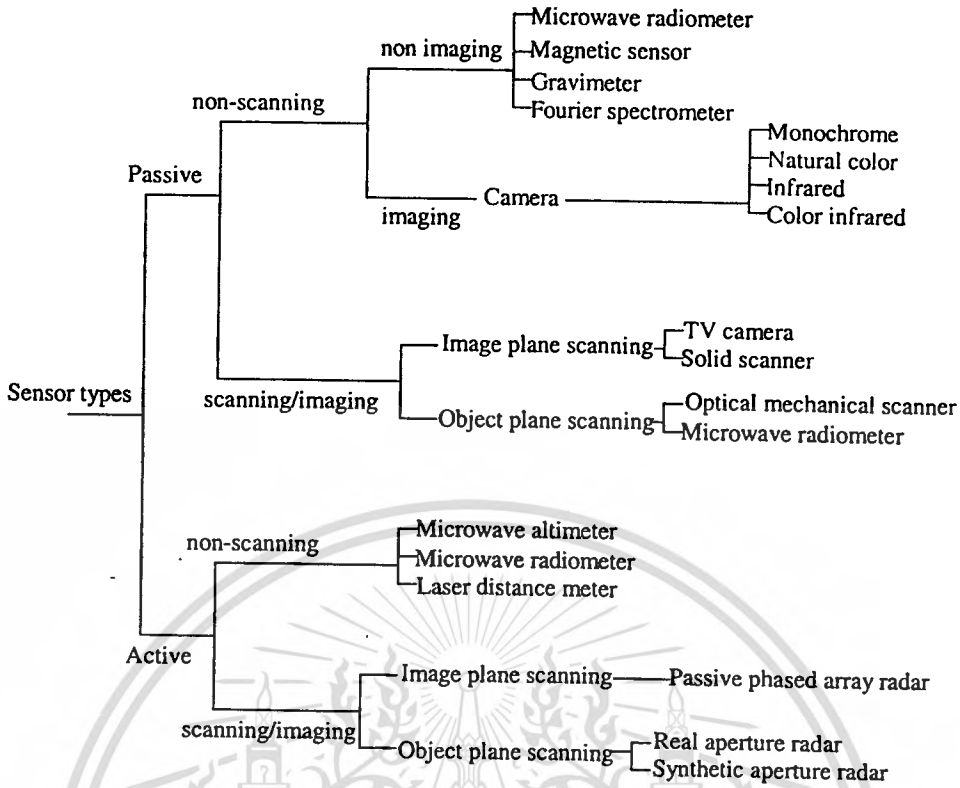


Figure 2.9 Sensor systems

It is the fact that the sun is a source of energy or radiation. The sun provides a main source of energy for remote sensing. The sun's energy is either reflected, as it is for visible wavelengths, or absorbed and then reemitted, as it is for thermal infrared wavelengths. Remote sensing systems which measure energy that is naturally available are called passive sensors. Passive sensors can only be used to detect energy when the naturally occurring energy is available. For all reflected energy, this can only take place during the time when the sun is illuminating the earth. There is no reflected energy available from the sun at nighttime. On the other hand, the energy that is naturally emitted by a target can be detected in day or night, as long as the amount of energy is large enough to be recorded.

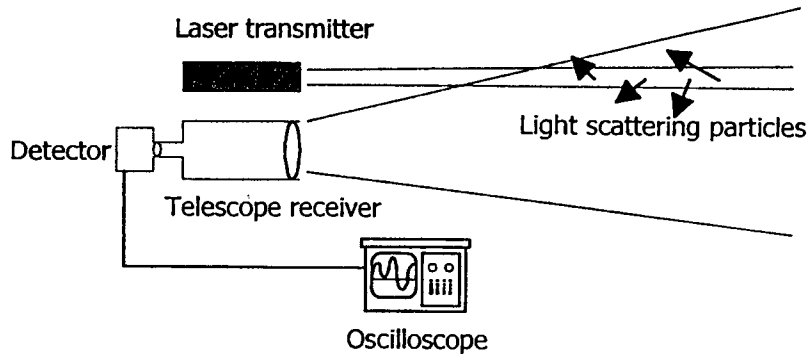
Active sensors, on the other hand, provide their own energy source for illumination. The sensor emits radiation which is directed toward a target to be investigated. The radiation reflected from that target is

detected and measured by the sensor. Advantages for active sensors include the ability to obtain measurements anytime, regardless of the time of day or season. Active sensors can be used for examining wavelengths that are not sufficiently provided by the sun, such as microwaves or millimeter wave. However, active systems require the generation of a fairly large amount of energy to adequately illuminate targets. Some examples of active sensors are a Millimeter Wave Cloud Radar (MWCR) and synthetic aperture radar (SAR).

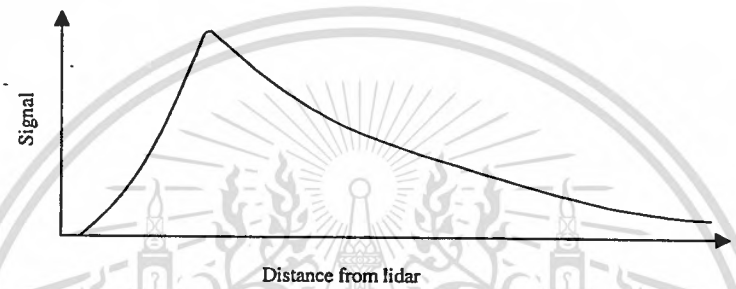
Another active sensor system, similar in some respects to radar, is lidar (light detection and ranging). A lidar transmits pulses of laser light, at various visible or near infrared wavelengths, to a target from which some of the light reflects. Travel times for the round-trip are the measured parameter. We can operate lidar instruments as profilers and as scanners, day and night. In addition, lidar can serve either as a ranging device to determine altitudes (topography mapping) or as a particle analyzer for the mixing gaseous molecules in the atmosphere.

2.7 Lidar System

Light is scattered and attenuated by molecules, aerosols, and cloud (water or ice) particles in the atmosphere. The sky can be clear and blue or hazy and white. Red sunsets are a beautiful manifestation of the scattering and attenuation of sunlight. Clouds can appear white, grey, or dark depending on conditions. Light scattering and attenuation can be used to investigate the atmosphere using a remote-sensing instrument called a lidar. A lidar system uses laser pulses to measure atmospheric constituents such as aerosol particles, ice crystals, water vapor, or trace gases (e.g. ozone). Profiles of these atmospheric components as a function of altitude or location are necessary for weather forecasting, climate modeling, and environmental monitoring.



(a) Basic component of lidar



(B) Lidar signal as function of range

Figure 2.10 Basic lidar system

A lidar transmits short pulses of laser light into the atmosphere. The laser beam loses light to scattering as it travels. At each range, some of the light is backscattered into a detector. In Fig. 2.10(a), because the light takes longer time to return from the more distant ranges, the time delay of the return pulses can be converted to the corresponding distance between the atmospheric scatterer and the lidar. The end result is a profile of atmospheric scattering versus distance. In Fig. 2.10(b), analysis of this signal can yield information about the distribution of aerosols in the atmosphere. The amount of backscatter indicates the density of the scatters. This can be used to measure cloud base height or track plumes of pollution. Other properties of the atmosphere can also be deduced from the lidar return signals such as frequency shifts, depolarization, and backscatter intensity from multiple wavelengths. Currently, there are

several types of lidar systems that have been used to study the atmosphere.

2.7.1 Scattering Lidar

Most of early laser radar or lidar systems are scattering lidar or Mie lidar. A commonly used technique is to detect Rayleigh and Mie backscatter from the atmospheric molecules and aerosols, respectively. This technique reveals aerosol distributions and optical properties of the atmospheric boundary layer. Several authors use this method to investigate the aerosol distributions, clouds and precipitation elements in the atmosphere [1], [2].

Rayleigh and Mie backscattering are considered as elastic scattering. Rayleigh scattering of light occurs in a gas when the wavelength of the incident light is remote from that of any absorption line while Mie scattering by atmospheric aerosol is known to be an important contributing factor to the scattered signal received from the atmosphere and is normally evaluated by a classical calculation based on electromagnetic scattering theory. The results depend critically on the shape and size of the particle.

2.7.2 Differential Absorption Lidar (DIAL)

The DIAL measures the atmosphere using two wavelengths: one on the absorption peak and another off the absorption peak of a selected atmospheric constituent, which is shown in Fig. 2.11. The independent wavelength of backscatter is smooth, while the absorption is highly selective wavelength. Thus, the absorption difference between backscattered signals on and off the absorption peak reveals the quantity of the absorbing media. Fredriksson and Hertz [3] used DIAL measurements of atmospheric NO_2 using mobile lidar system. Browell [4] used an aircraft-based DIAL to map ozone levels in the lower troposphere. DIAL is particularly valuable for urban and industrial monitoring as well

as analysis of gases from geogenic origin. Tropospheric pollution monitoring is most frequently done using differential absorption lidar.

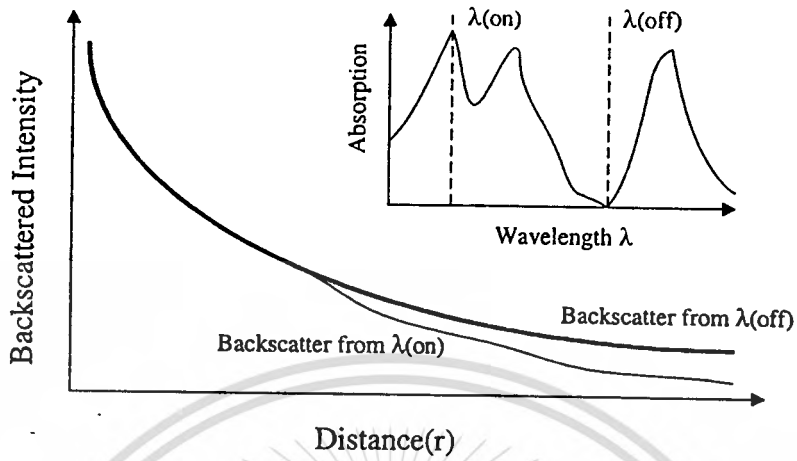


Figure 2.11 Illustrates the DIAL principle

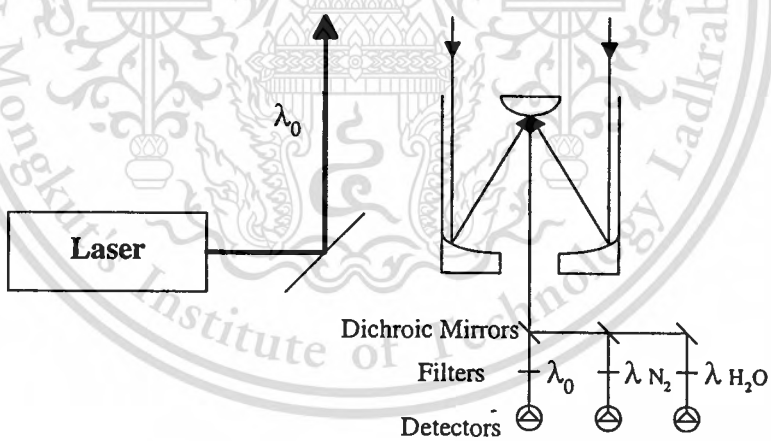
2.7.3 Rayleigh-Doppler Lidar

Rayleigh-Doppler lidar transmits pulses laser light into the atmosphere and receives the backscattering light occurred by atmospheric molecules. The system observes the atmospheric density and temperature profiles from the received light intensity. Rayleigh Doppler lidar is used for observing atmospheric winds by measuring the spectrum of reflected light. The frequency shift, which can estimate the molecular velocity, between the original pulsed laser light and the received light is caused by the Doppler shift of the motions of atmospheric molecules. This system is uniquely able to measure the wind of neutral atmosphere since it employs Rayleigh scattering by atmospheric molecules.

2.7.4 Raman Lidar

The Raman lidar is laser remote sensing instrument that measures vertical profiles of water-vapor mixing ratio, several cloud and aerosols. Figure 2.12, the Raman lidar has independent measurement channels that record range-resolved backscatter signals from molecular water vapor, molecular nitrogen, and combined Rayleigh and aerosol

contributions. Primary quantities obtained from these backscatter signals are range-resolved vertical profiles of water-vapor mixing ratio, aerosol scattering ratio, and backscatter depolarization ratio. Raman lidar detect selected species by monitoring the wavelength-shifted molecular return produced by vibrational Raman scattering from the chosen molecule. By taking the ratio of the signal at the water-vapor wavelength to the signal at the nitrogen wavelength, most of the range-dependent terms drop out, and one is left with a quantity that is almost directly proportional to the water-vapor mixing ratio expressed as grams of water vapor per kilogram of dry air. Similarly, by taking the ratio of the signal at the laser wavelength to the signal at the nitrogen wavelength, one is left with the aerosol backscatter ratio. Finally, analysis of the polarization dependence of the backscatter signal at the laser wavelength provides information on particle shape; spherical particles, cloud droplets, do not depolarize the laser backscatter, whereas nonspherical particles such as ice crystals in cirrus clouds can significantly depolarize the laser backscatter.



Water-vapor mixing ratio α $S(H_2O)/S(N_2)$

Aerosol backscatter ratio α $S(O)/S(N_2)$

Figure 2.12 Raman lidar

2.7.5 Multiwavelength lidar

Lidar backscatter varies as a function of wavelength, as well as atmospheric composition and aerosol properties. By utilizing lidars at different wavelengths we have investigated measurement techniques for differentiating between the hygroscopic and non-hygroscopic components of the aerosol. Hygroscopic aerosol particles swell and increase in size quite dramatically at relative humidity above about 80% whereas their non-hygroscopic counterparts are unaffected by relative humidity. Different lidar wavelengths respond to this size change to varying degrees and we have utilized this differential to retrieve a characteristic size and number of these hygrosocopic particles. This knowledge will allow us to retrieve cloud condensation nucleus spectra remotely and significantly enhance our ability to monitor aerosol-cloud climate interactions.

2.7.6 Fluorosence Lidar

Fluorescence lidar illuminates the target with a specific wavelength of radiation and is capable of detecting multiple wavelengths of fluoresced radiation. This technique provides another powerful tool for air pollution studies. It is based on the detection of organic, fluorescent material from the atmosphere. Resonance scattering or resonance-fluorescence can occur only when the frequency of the incident radiation coincides with an absorption line or band of the molecules. The technique can be used for long-range identification of biological particulates and cyclic aromatic hydrocarbon derivatives. This has potential applications for monitoring environmental pollution, industrial waste emission. Granatstein, Rhinewine and Fitch [5] used a fluorescence lidar to study the resonant absorption of CH_4, CO_2 from a steam plume source.

Chapter 3

Tropospheric Aerosol Observation

3.1 Introduction

After the laser was invented, the various lasers have been applied in many field, and in the field of work for detecting the target that is well known as radar or remote sensing. In this field, the laser have been applied to perform in several platforms such as space borne, air borne, or ground-based radar. The laser has been used to study the composition of many mixing particles in the atmosphere including the aerosol distributions. A system is used to probe the atmosphere is called laser radar or optical remote sensing or lidar (light detection and ranging). Currently, there are several lidar systems that are used in this filed depending on the purpose of measurements, such as the measurement of ozone, vapor water, sodium and etc., including the aerosol measurement in boundary layer. Each behavior of measurements must use specific technique that has been described in the chapter 2.

It is obvious that the earth is covered with various mixing gasses, which are divided into several layers with respect to the characteristics of the atmospheric temperature. The lowest atmospheric layer is called troposphere in which many phenomena such as rain, snow, cloud, fog or aerosol only occurred. Moreover, the boundary layer is the region in the troposphere, i. e., only 3 km above the ground surface. Much pollution that was made by human activity occurs in this region. The laser was used to study the composition of mixing particle including the pollutant or aerosol distributions in the atmosphere. In this thesis, the ground-based laser radar is used to measure the aerosol distributions in the vertical axis about 3 km height based on elastic light scattering techniques.

3.2 Aerosol

What are aerosols and why are they so important? Aerosols are tiny liquid and solid particles suspended in the air. Most occur naturally, originating from volcanoes, dust storms, forest and grassland fires and living vegetation. Human activities also generate aerosols such as the burning of fossil fuels. Aerosols tend to cause cooling of the earth's surface immediately below them. Because most aerosols reflect incoming sunlight back into space, they have a direct cooling effect by reducing the amount of solar radiation that reaches the surface. The magnitude of this cooling effect depends on the size and composition of the aerosol particles, as well as the reflective properties of the underlying surface. It is thought that aerosol cooling may partially offset expected global warming that is attributed to increases in the amount of carbon dioxide from human activity. Aerosols are also an indirect effect on climate by changing the properties of clouds. Changing aerosols in the atmosphere can change the frequency of cloud occurrence, cloud thickness, and rainfall amounts. Both effects increase the amount of sunlight that is reflected to space without reaching the surface.

There are mainly three types of aerosols significantly affect the earth's climate. The first is the volcanic aerosol layer which forms in the stratosphere after major volcanic eruptions like Mt. Pinatubo. The dominant aerosol layer is actually formed by sulfur dioxide gas which is converted to droplets of sulfuric acid in the stratosphere over the course of a week to several months after the eruption. They reflect sunlight, reducing the amount of energy reaching the lower atmosphere and the earth's surface, cooling them. The second type of aerosol that may have a significant effect on climate is desert dust which is composed of minerals. These particles absorb sunlight as well as scatter it. Through absorption of sunlight, the dust particles warm the layer of the atmosphere where they reside. The third type of aerosol comes from human activities. While a large fraction of human-made aerosols come from burning tropical forests, the major component comes in the form of sulfate aerosols created

by the burning of oil. The concentration of human-made sulfate aerosols in the atmosphere has grown rapidly since the start of the industrial revolution. The sulfate aerosols absorb no sunlight but they reflect it, thereby reducing the amount of sunlight reaching the earth's surface.

Aerosol measurements can also be used to study how the earth's atmosphere moves. Aerosols have been used to study the dynamics of the polar regions and the exchange of air between the troposphere and stratosphere. Currently, space-based instruments provide a way of making, at a reasonable cost, continuous, global observations of aerosols to help scientists measure how these particles influence climate. An example is the data from the SAGE II satellite instrument on NASA's Earth Radiation Budget Satellite (ERBS) [6] have shown that after the Mt. Pinatubo eruption, the optical depth of the stratospheric aerosol increased up to 100 times in certain locations. Optical depth is a general measure of the capacity of a region of the atmosphere to prevent the passage of visible light through it.

3.3 Lidar equation

Atmospheric aerosols are influence in the radiative forcing of the earth's climate, as they influence the radiation balance of the Earth, mostly through scattering and absorption processes. Most aerosols are found in the lower troposphere and contribute significantly to the haze often visible during early morning hours. These aerosols have residence of few days, and thus are not distributed homogeneously in the atmosphere. Currently, there are still many uncertainties concerning the spatial distribution, the shape and chemical composition of the tropospheric aerosols, and especially the chemical coupling between particulate matter and ozone. Extensive studies of the atmospheric aerosol in the troposphere have been conducted by using the lidar techniques in many laboratories. Based on the elastic backscattering of emitted laser radiation by the atmospheric aerosols, the lidar techniques

allow the real-time visualization of the suspended particles in the atmosphere with high temporal and spatial resolution.

Figure 3.1 Lidar system

In Fig. 3.1, when the laser energy is transmitted in vertical axis, the echo signal, which is inversely proportional with square of distance, can be obtained from the backscattering of the laser light. The backscattering can be separated in two causes. One is the laser light scattering only from the mixing gaseous molecules which know as Rayliegh scattering. The second is the scattered light occurring from dust, aerosol, and the whole particles in the laser beam. Usually, the aerosol diameter is larger than the laser wavelength, for this reason this scattering is called Mie scattering.

The echo signal may be quantitatively interpreted in terms of the lidar range equation. For a vertically pointing, the instantaneous

received power, p(z), due to backscattering from height z can be expressed in the form

$$p(z) = p_0 \eta(z) \frac{c\tau}{2} \cdot \frac{A_e}{z^2} B(z) T^2(z), \tag{3.1}$$

where

p_0 = transmitted laser power, (W)

$\eta(z)$ = optical factor

c = speed of light, (ms^{-1})

τ = transmitted laser pulse length, (s)

A_e = receiver aperture, (m^2)

z = altitude from ground surface in vertical axis, (m)

$B(z)$ = unit volume backscattering coefficient, ($m^{-1}sr^{-1}$)

$T(z)$ = one way atmospheric transmittance

For ranges beyond complete overlap between the transmitter and receiver beam path, the lidar equation can be written as

$$P(z) = \frac{CEB(z)T^2(z)}{z^2} \tag{3.2}$$

where C is the lidar system calibration constant which includes the losses in the transmitting and receiving optics, the effective receiver aperture and E is the transmitted pulse energy($p_0\tau$). The transmission factor $T(z)$ follows the exponential law of attenuation with the Beer-Lamber law[1], [2]. It is related to the integrated extinction by

$$T(z) = \exp(-\int_0^z \sigma(z)dz) \tag{3.3}$$

where $\sigma(z)$ is the atmospheric unit volume extinction coefficient(m^{-1}) along the path z. For two-way atmospheric transmittance between lidar and z is

$$T^2(z) = \exp(-2\int_0^z \sigma(z)dz) \tag{3.4}$$

or

$$\sigma(z) = -\frac{1}{2T^2(z)}\frac{dT^2(z)}{dz} \tag{3.5}$$

If S is defined as the ratio between the extinction coefficients σ(z) and backscattering coefficients $B(z)$, Eq. (3.5) becomes

$$B(z) = -\frac{1}{2ST^2(z)}\frac{dT^2(z)}{dz} \tag{3.6}$$

The lidar equation then is

$$P(z) = -\frac{EC}{2Sz^2}\frac{dT^2(z)}{dz} \tag{3.7}$$

The transmittance can be obtained by integrating from height 0 to z

$$T^2(z) = 1 - \frac{2S}{CE}\int_0^z P(z)z^2 dz \tag{3.8}$$

As noted earlier, the constituents in the atmosphere include the gaseous air molecules and aerosol particles. The molecular and aerosol scattering contributions can be treated by superposition so that $B(z)$ may be resolved as

$$B(z) = B_m(z) + B_a(z) \tag{3.9}$$

where $B_a(z)$ is the aerosol backscattering coefficients and $B_m(z)$ is the Rayleigh backscattering coefficients. So that

$$T^2(z) = T_m^2(z) T_a^2(z) \tag{3.10}$$

$$\sigma_m(z) = -\frac{1}{2T_m^{\,2}(z)} \frac{dT_m^{\,2}(z)}{dz} \tag{3.11}$$

$$B_m(z) = -\frac{1}{2S_m T_m^{\,2}(z)} \frac{dT_m^{\,2}(z)}{dz} \tag{3.12}$$

where $B_m(z)$ and $\sigma_m(z)$ are the backscattering coefficients and the corresponding extinction coefficients for the molecular in the atmosphere, respectively. For the lidar measurements in here the lidar inversion technique was applied [1]. In this technique, it is generally assumed that the Rayleigh extinction and backscattering coefficients are known functions of height and are obtained by the use of an atmospheric model. Besides, the aerosol backscatter-to-extinction ratio S_m is assumed to be constant with height. Similarly, $B_a(z)$ and $\sigma_a(z)$ are the backscattering coefficients and the corresponding extinction coefficients for the aerosol particles in the atmosphere, respectively.

$$\sigma_a(z) = -\frac{1}{2T_a^{\,2}(z)} \frac{dT_a^{\,2}(z)}{dz} \tag{3.13}$$

$$B_a(z) = -\frac{1}{2S_a T_a^{\,2}(z)} \frac{dT_a^{\,2}(z)}{dz} \tag{3.14}$$

The lidar equation now becomes

$$P(z) = \frac{CE[B_m(z) + B_a(z)] T_m^{\,2}(z) T_a^2(z)}{z^2} \tag{3.15}$$

If $B_a(z)$ is expressed in terms of the two-way transmittance by applying Eq. (3.15), then becomes

$$P(z)=\frac{CE\left[B_m(z)+-\frac{1}{2S_aT_a^{\,2}(z)}\frac{dT_a^{\,2}(z)}{dz}\right]T_m^{\,2}(z)T_a^2(z)}{z^2} \tag{3.16}$$

This can be written in the from

$$\frac{dT_a^2}{dz}=2S_aB_m(z)T_a^2(z)-\frac{2S_aP(z)z^2}{ECT_m^2(z)} \tag{3.17}$$

and when solved for $T_a^2(z)$, it is

$$T_a^2(z)=\exp(2S_a\int_0^z B_m(z)dz)\left[1-\frac{2S_a}{EC}\int_0^z\frac{P(z)z^2}{T_m^2(z)}\exp(-2S_a\int_0^z B_m(z)dz)dz\right] \tag{3.18}$$

From Eq. (3.14) the two-way transmittance are in the form

$$T_a^2(z)=T_m^{-2S_a/S_m}\left[1-\frac{2S_a}{EC}\int_0^z\frac{P(z)z^2}{T_m^2(z)}T_m^{2S_a/S_m}(z)dz\right] \tag{3.19}$$

Substituting this back into (3.15) and solving for $B_a(z)$ then yields

$$B_a(z)=\frac{P(z)z^2\exp(-2(S_a-S_m)\int_o^z B_m(z)dz)}{EC-2S_a\int_0^z P(z)z^2\exp(-2(S_a-S_m)\int_0^z B_m(z')dz')dz}-B_m(z) \tag{3.20}$$

The lidar can be calibrated by solving Eq. (3.15) for CE at the range z*

$$CE = \frac{P(z^*)z^{*2}}{[B_m(z^*) + B_a(z^*)]T_m^{\ 2}(z^*)T_a^2(z^*)} \tag{3.21}$$

Generally, the data analysis is suitable form in the backscattering ratio of Mie and Rayliegh backscattering cross section. It is expressed as

$$S(z) = \frac{B_m(z) + B_a(z)}{B_m(z)} \tag{3.22}$$

If any region in the laser beam which disappear aerosols, the $B_a(z)$ approaches to zero so that the scattering ratio is always equal to one. On the contrary, the scattering ratio is more than one, when the aerosol particles occur in concerning region.

3.4 System Configuration

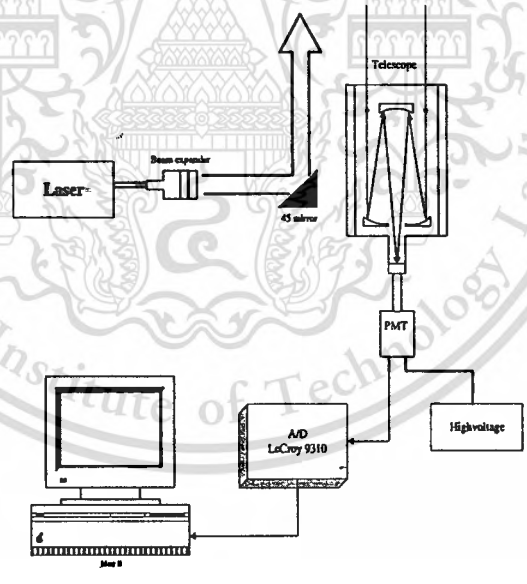


Figure 3.2 System configuration

The configuration of laser radar system is shown in Fig. 3.2. A transmitter is the Nd:YAG laser, which is shown in Fig. 3.3, that emits the secondary harmonic wavelength (532 nm.), 25 mJ pulse energy, 20 ns

pulse duration. Laser light is transmitted in the vertical axis with respect to the ground surface. A telescope with 28 cm diameter, in Fig. 3.4, is used as the receiver. The telescope collects the scattering laser light from molecules and particles in the field of view of the telescope. The collected lights pass through a narrow band filter and then focusing on Hamamatsu photo multiplier that converts the light energy to electrical signal. A 9304 LeCroy digital oscilloscope performs the signal acquisition, which communicates to a personal computer. One recorded signal is the total average of 500 repetitions with the 150 m height resolution. Finally, the digital data are transferred to a personal computer where the signals are processed and analyzed. The detailed system is shown in Table 3.1.

Figure 3.3 Photograph of Nd: YAG laser

Figure 3.4 Photograph of the receiver

Table 3.1 Characteristics of lidar system

Transmitter:	
Laser	Nd: YAG, SHG(532 nm)
Output	180 mJ
Repetition	10 Hz
Beam div.	0.1 mrad
Pulse width	20 nsec
Receiver:	
Telescope	Schmidt-Cassegrain 28 cm diameter
Detector	PMT, S-1 R3236 Hamamatsu

3.5 Experimental results

Having observed, Figure 3.5 is the photograph from oscilloscope at the output of photo multiplier tube, and the signal power are rapidly decreasing when the distance is increased. Moreover, the peak of return signal means that the altitude, about 300 m, of the over lap between laser beam and field of view of receiving telescope. The 500 signals are accomplished and analyzed to obtain the backscattering cross section, in Figure 3.6.

Figure 3.5 Photograph of the return signal from oscilloscope

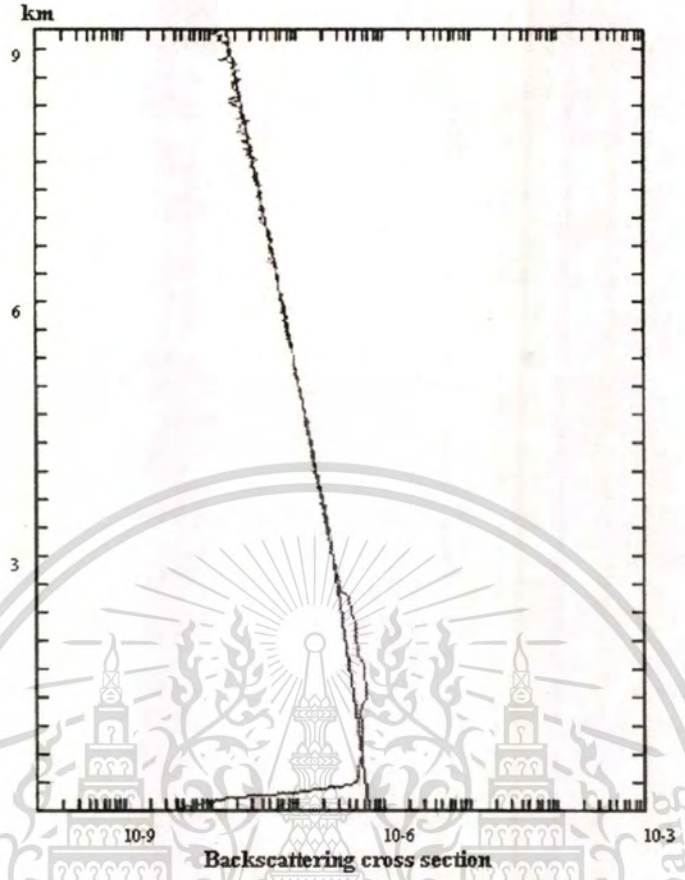


Figure 3.6 Backscattering cross section

Figure 3.6 shows the relation between the backscattering cross section in log scale with altitude while a straight line represents the Rayleigh backscattering cross section and the thick line is the backscattering cross section which is obtained from lidar data. The backscattering ratio is obtained in Fig. 3.7. Therefore, if the aerosol appears at any altitude, the scattering ratio will increase. When the observations are continuous, the scattering ratio profile in vertical axis is obtained in Fig. 3.8. This profile shows the located aerosols that vary in time and space.

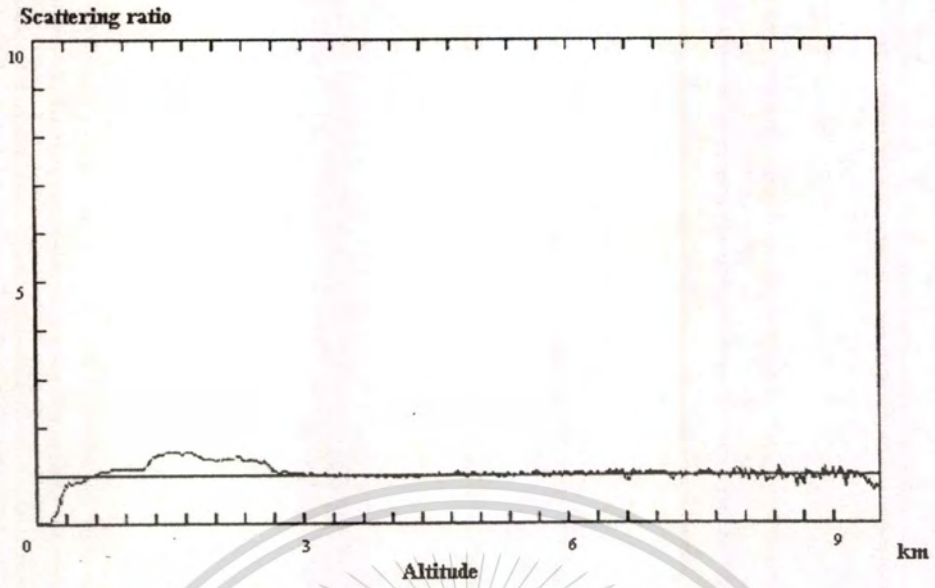


Figure 3.7 Scattering ratio

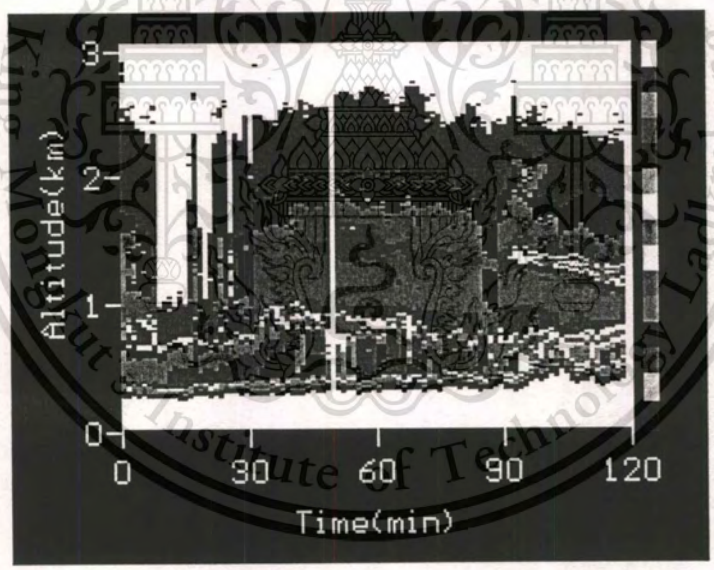


Figure 3.8 Aerosol profile

Chapter 4

Data Classification

4.1. Introduction

When a clustering technique is considered, which is one of major problems in the fields of pattern classification and recognition, there are many types of cluster analyses. The clustering algorithms require no training data and the main objective of its ability is to partition the data set of given multidimensional vectors into some groups (clusters) based on the certain degree of similarity. Several useful clustering algorithms have been collecteded by Duda and Hart [16], Jain and Dubes [11], Ton and Gonzlez [10]. As any well known and widely used clustering techniques, the K-means algorithm, the Hard *c*-mean (HCM) algorithm, self-organizing feature maps, competitive-learning network and adaptive resonance theory (ART) are considered. Each approach may give the different result on data partitioning depending on the specific conditions for clustering. One of the major problems encountered during the clustering of the given data is that the number of clusters is not known a priori. Normally, the approach to the determination of the cluster number is the full search of all cluster numbers and the selection of the optimum cluster set. The selection is usually hard to be realized without a priori knowledge. This problem has been known as the cluster validity problem. Accordingly, it has required a new method to estimate the optimum number of clusters.

4.2 Clustering Algorithms

In this section, two clustering algorithms are considered, i.e., Hard *c*-mean (HCM) and self-organizing clustering (SOC). They are employed into our proposed method. Let us consider the set of n vectors in the m-

dimensional feature space $X=\{\mathbf{x}_1, \mathbf{x}_2,..., \mathbf{x}_n\} \subset R^m$. These vectors are assumed to be clustered into c groups. Define the family of cluster sets $\{A_i, i=1, 2,..., c\}$ for X where the following equations are satisfied:

$$\bigcup_{i=1}^{c} A_i = X \tag{4.1}$$

$$A_i \cap A_j = \phi \quad \text{all } i \neq j \tag{4.2}$$

$$\phi \subset A_i \subset X \quad \text{all } i \tag{4.3}$$

The number of A_i is obviously given in $[2,n-1]$. Let us define that c^* is the number of the optimal clusters that can be obtained from the cluster validity criteria.

4.2.1 HCM Algorithm

HCM is a clustering technique by which a data set is partitioned into disjoint subgroups. It means that each data point $\mathbf{x}_j$ ($j=1, 2,..., n$) is assigned to only one data cluster. On this sense, these clusters are called partitions. At the beginning process, the number of c is given and the initial cluster center $\mathbf{z}_1, \mathbf{z}_2,..., \mathbf{z}_c$ are chosen where $\mathbf{z}_i=\{z_{i1}, z_{i2},..., z_{im}\}$ ($i=1, 2,..., c$). After the input vector is fed into all clusters, the cluster is selected based on the similarity or the distance measurement. The selected cluster is expressed with the selection index

$$\delta_{ij} = \begin{cases} 1 & \mathbf{x}_j \in A_i \\ 0 & \mathbf{x}_j \notin A_i \end{cases}. \tag{4.4}$$

For the hard partition, $\bigvee_{i=1}^{c} \delta_{ij} = 1$ and $\delta_{ij} \wedge \delta_{kj} = 0$ should be satisfied for all j. Defined a matrix U which comprises of δ_{ij} ($i=1, 2,..., c; j=1, 2,..., n$). We then define a hard c-partition space as the following matrix set:

$$M_C = \{U | \delta_{ij} \in \{0,1\}, \sum_{i=1}^{c} \delta_{ij} = 1, 0 < \sum_{j=1}^{n} \delta_{ij} < n \} \quad (4.5)$$

The HCM objective function is given by

$$J(U, z) = \sum_{i=1}^{c} \sum_{j=1}^{n} \delta_{ij} d_{ij}^2 \quad (4.6)$$

where U is partition matrix, the parameter z is a vector of cluster centers, d_{ij} is a Euclidean distance between the j^{th} data point and i^{th} cluster center z_i,

$$d_{ij} = \left\| x_j - z_i \right\| = \left[\sum_{k=1}^{m} \left(x_{jk} - z_{ik} \right)^2 \right]^{1/2} \quad (4.7)$$

and $z_{ik}(i=1, 2,..., c, k=1, 2,..., m)$ is

$$z_{ik} = \frac{\sum_{j=1}^{n} \delta_{ij} x_{ik}}{\sum_{j=1}^{n} \delta_{ij}} \quad (4.8)$$

The optimum partition matrix, i.e., U*, is obtained by using the following condition,

$$J(U^*, z^*) = \min_{U \in M_c} J(U, z) \quad (4.9)$$

4.2.2 SOC Algorithm

According to the conventional SOC algorithms [11], [12], [13], the system estimates the cluster distribution based on the Euclidean distance.

There are c clusters with the cluster pattern $\mathbf{m}_i(n)$, an error variance $\mathbf{V}_i(n)$ and the number of members $t_i(n)$. For the input vector, a similarity is first calculated. The suitable cluster is then selected. Once the input sample is determined belong to the suitable cluster, the information of this selected cluster is updated. The information of the cluster is evaluated as

$$u_i(n) = \sum_{j=1}^{n} \delta_{ij} \tag{4.10a}$$

$$t_i(n) = \sum_{j=1}^{n} \Gamma^{u_i(n)-u_i(j)} \delta_{ij} \tag{4.10b}$$

$$\mathbf{m}_i(n) = \frac{1}{t_i(n)} \sum_{j=1}^{n} \Gamma^{u_i(n)-u_i(j)} \delta_{ij} \mathbf{x}_j \tag{4.10c}$$

Since a forgetting factor Γ is defined as $0<\Gamma<1$ and $\Gamma \approx 1$, every node deals with the new data as more important than previous data. This forgetting factor is useful for an adaptation mechanism. In addition, the variance of a cluster is also calculated by

$$\mathbf{V}_i(n) = \frac{1}{t_i(n)} \left[\sum_{j=1}^{n} \Gamma^{u_i(n)-u_i(j)} \delta_{ij}(j) \mathbf{e}(j|n) \mathbf{e}(j|n)^T \right] \tag{4.11}$$

where the error vector $\mathbf{e}(j|n)$ is defined as

$$\mathbf{e}(j|n) = \mathbf{x}_j - \mathbf{m}_i(n) \tag{4.12}$$

and each element is assumed to be independent of others.

Note that $\mathbf{V}_i(n)$ is assumed to be a diagonal matrix. From the above vectors and a matrix, the similarity is defined as

$$d_i(\mathbf{x}_n, n-1) = \exp\left[-\frac{\alpha}{2m} \mathbf{e}_i(n|n-1)^T \mathbf{V}_i^{-1}(n\text{-}1) \mathbf{e}_i(n|n-1) \right] \qquad (4.13)$$

where α is a rectification factor defined in [25].

4.3 Cluster Validity Criterion

Cluster validity is an important issue which is crucial for the practical application on clustering. At present, there have been some existing validity criteria such as partition coefficient, abbreviated as PC, and partition entropy, i.e., PE, which are proposed by Pal and Bezdek [17] are two simple cluster validity criteria associated with a fuzzy c-mean algorithm (FCM). Other criteria have been proposed by Xie and Bini, i.e., XB [18], Fukuyama and Sugeno, i.e., FS [19]. The analysis on these criteria have been presented by Pal and Bezdek [17]. Recently, another approach based on the structural characteristics around the optimal cluster number in the partitioning process has been presented by Kim and Park, i.e., SV [20].

Since most of clustering methods need to assume the number c of clusters as a priori information, a validity criterion becomes an important issue in order to find an optimal c^* cluster. In this section, we consider an algorithm, that is a total fitting level method (TFL), for the cluster validity criterion. In this study, we assume that the distribution of observed data is Gaussian. The TFL is based on the difference between a histogram which is calculated from the i^{th} cluster member and the normal probability distribution function abbreviated as p.d.f. of that cluster. Figure 4.1 shows simple input data that consists of three compact classes. The optimal cluster number is three, i.e., c^*=3. Each cluster has the probability distribution and the histogram.

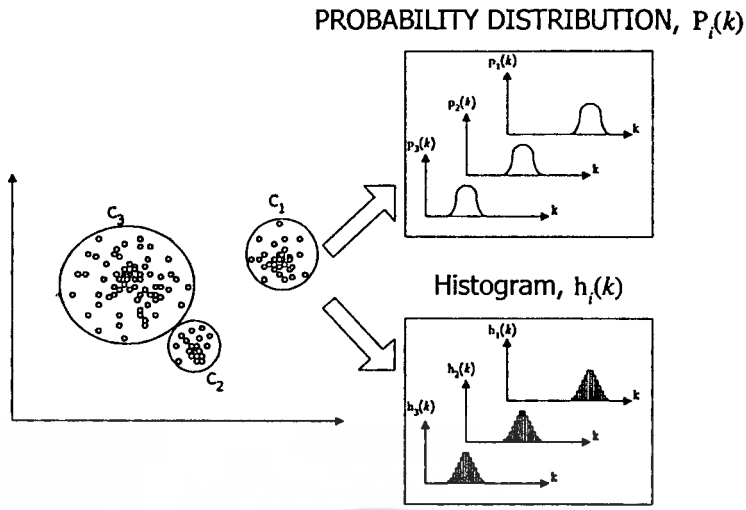


Figure 4.1 Example of data

We define that

$h_i(k)$ is the histogram of the i^{th} cluster and

$p_i(k)$ is the p.d.f. of the i^{th} cluster with mean $\mathbf{x}_i$ and variance V_i.

If a clustering algorithm does not provide mean and variance, they can be estimated from vectors while belong to this cluster as

$$\hat{\mathbf{x}}_i = \frac{\sum_{j=1}^{n} \delta_{ij} \mathbf{x}_j}{\sum_{j=1}^{n} \delta_{ij}} \tag{4.14}$$

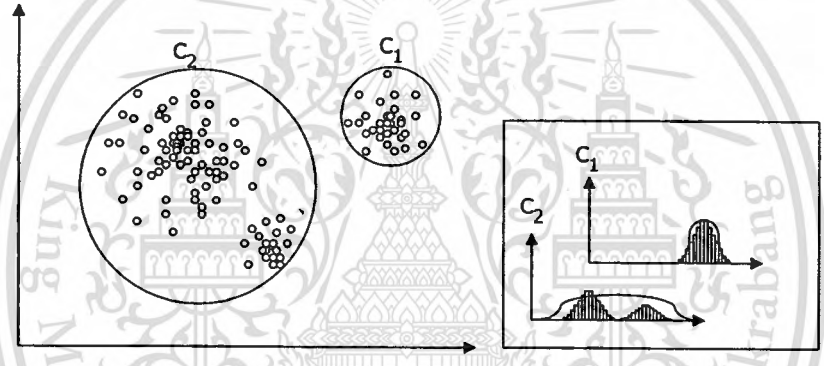
$$\mathbf{V}_i = \left[\frac{\sum_{j=1}^{n} \delta_{ij} \left\| \mathbf{x}_j - \mathbf{m}_i \right\|^2}{\sum_{j=1}^{n} \delta_{ij}} \right]^{1/2} \tag{4.15}$$

where δ_{ij} is the index of selection in (4.4) and $\mathbf{m}_i$ is the cluster pattern at the i^{th} cluster. The fitting level at the i^{th} cluster is defined as

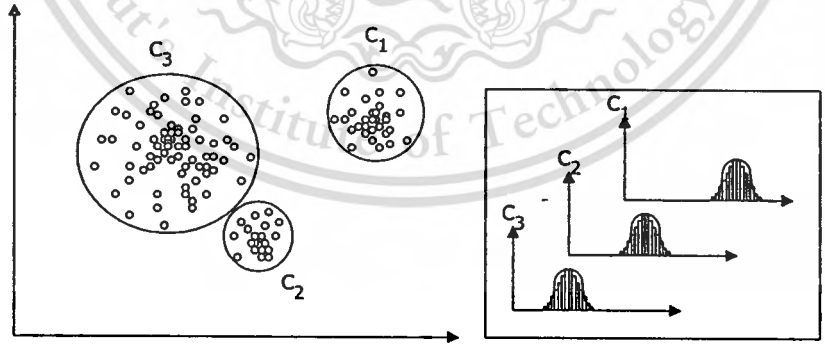
$$FL_i = \sum_{k \varepsilon c_i} |p_i(k) - h_i(k)| \tag{4.16}$$

Figure 4.2 shows the examples. From statistical property, i.e., $\sum_{k=-\infty}^{\infty} p_i(k) = 1$ and $\sum_{k=-\infty}^{\infty} h_i(k) = 1$, the fitting level lies between 0 to 2, i.e., $0 < FL_i < 2$. The total fitting level (TFL) is the average of the fitting level at every cluster,

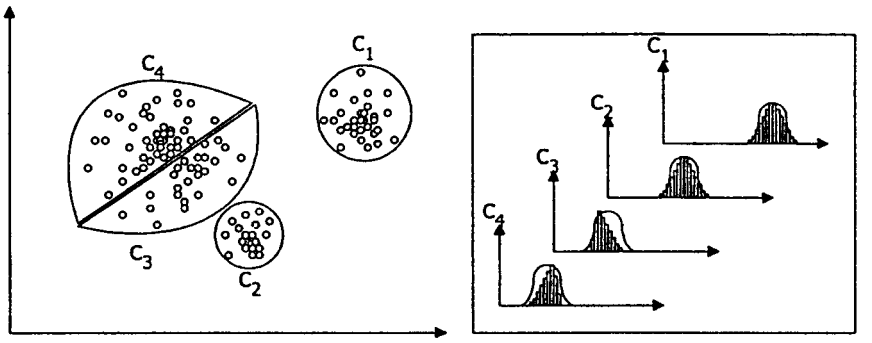
$$TFL(c) = \frac{1}{c} \sum_{i=1}^{c} FL_i \tag{4.17}$$



(a) Under partition($c<c^*$)



(b) Optimal partition($c=c^*$)



(c) Over partition($c>c^*$)

Figure 4.2 Three states of clustering

In Fig. 4.2(a), there are two clusters (c=2). This is the under partition. Figure 4.2(b) has three compact clusters (c=3). The FL of compact partition is away less than the under partition. Ideally, it closes to zero at compact partition. Moreover, in the case of over partition in Fig. 4.2(c), the FL increases again. Then the optimal cluster number c^* is obtained by solving $\min_{2\le c\le n-1} \mathrm{TFL}(c)$. Figure 4.3 shows one of example.

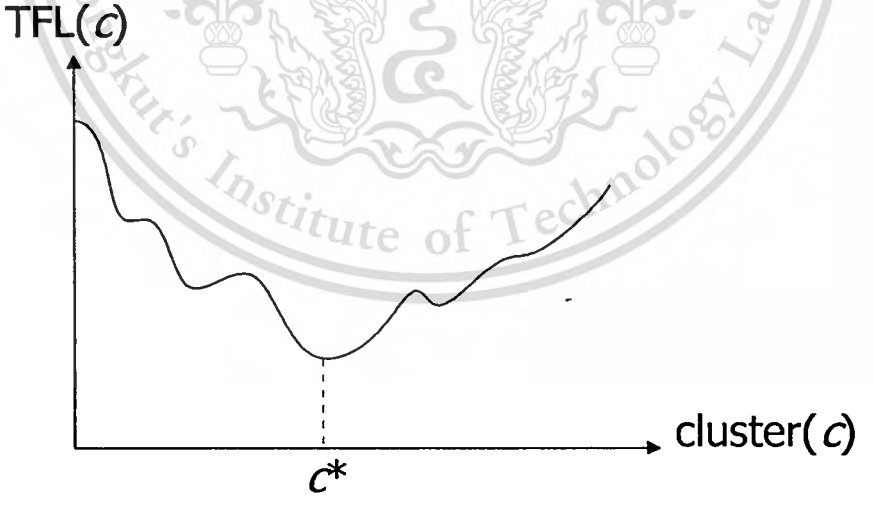


Figure 4.3 The validity index function TFL(c)

4.3.1 Clustering with TFL Algorithm

In this section, we apply the TFL with HCM (TFL-HCM) and SOC (TFL-SOC). To find the optimal number of clusters, the system starts with two clusters and increases until the enough number c_{max}. After that, we obtain the optimal number of clusters by solving the TFL algorithm. The following steps of process are

Step 0

Given input vectors $\{x_1, x_2, \ldots\ldots, x_n\}$ and set c=2.

Step 1

For TFL-HCM, initialize the U matrix; $U^{(0)} \in M_c$.

For TFL-SOC, give the random initial value of the cluster information

Step 2

For TFL-HCM, compute the cluster center z_{ik} in (4.8) and update $U^{(L)} \rightarrow U^{(L+1)}$ with the index of selection in (4.5).

For TFL-SOC, compute and update the cluster information in (4.10).

Step 3

For TFL-HCM, if $\| U^{(L+1)} - U^{(L)} \| < \varepsilon$ or

For TFL-SOC, if $\| m_i^{(L+1)} - m_i^{(L)} \| < \varepsilon$, go to the next step: Otherwise set L=L+1 and return to step 2. The ε is a tolerance level.

Step 4

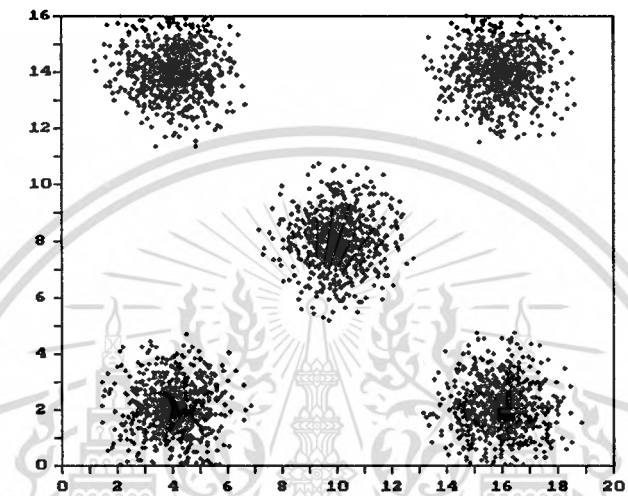
For TFL-HCM and TFL-SOC, compute the fitting level (FL) and the total fitting level (TFL) in (4.16) and (4.17), respectively.

Step 5

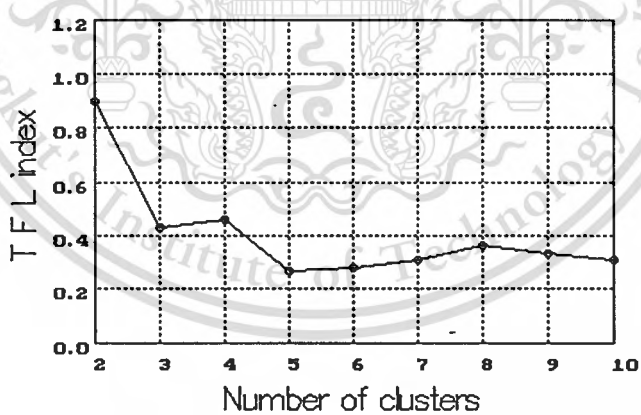
If $c < c_{max} < n$, set c=c+1, and then go to step 1. Otherwise find the optimal number of clusters c^* with min TFL(c). And then stop the whole processing.

4.4 Experimental Results

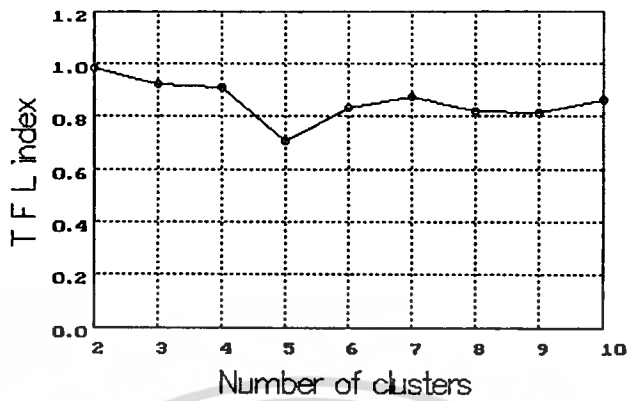
The simple data set is generated for the test on this algorithm. Figure 4.4(a) shows the five simple separated clusters. Each cluster has the normal distribution. Every cluster runs with 500 data points. The results of TFL-HCM and TFL-SOC algorithms show the same optimal cluster number (c^*=5) in the Fig.4.4(b) and Fig.4.4(c), respectively.



(a) Data S_1



(b) TFL-HCM validity function



(c) TFL-SOC validity function

Figure 4.4 Simple data s_1

Three complex data are generated and analyzed by the proposed algorithm. Figure 4.5-4.7 show the given data sets.

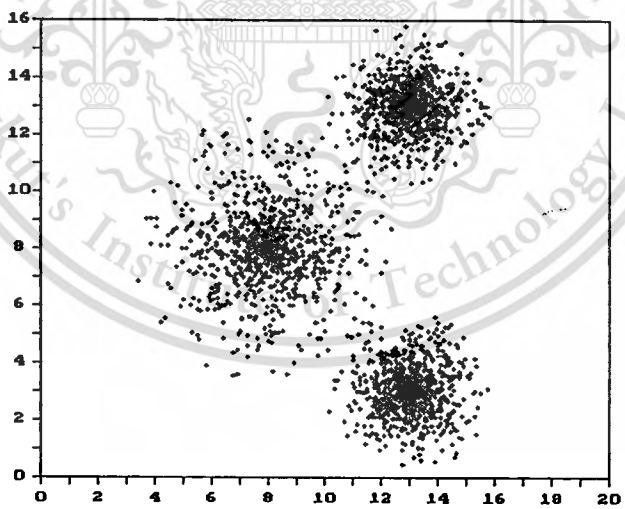


Figure 4.5 Data set C_1

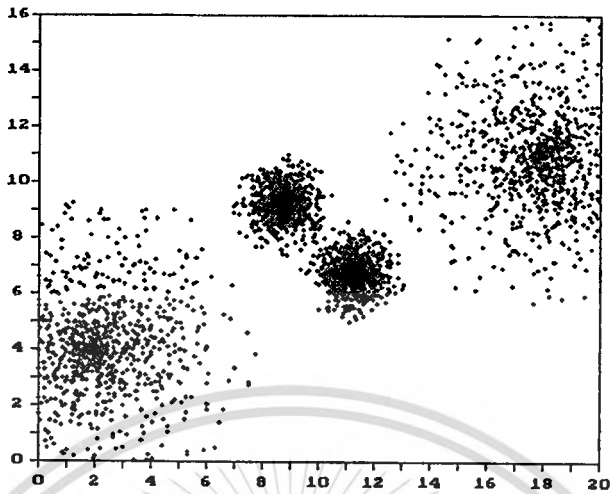


Figure 4.6 Data set C_2

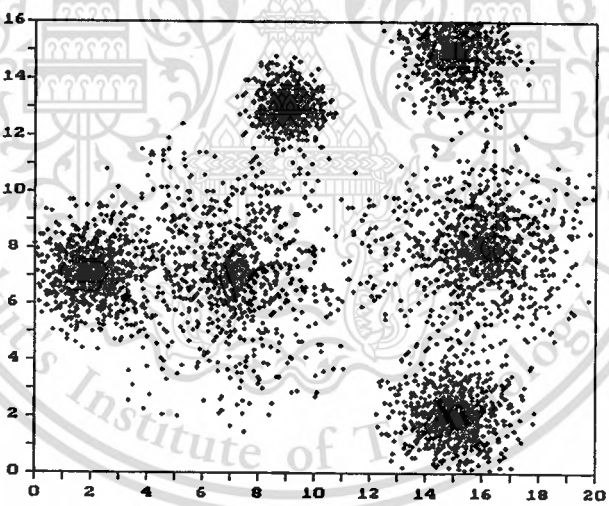


Figure 4.7 Data set C_3

Table 4.1 Validity index of data C_1

c	TFL-SOC	TFL-HCM
2	0.873275	0.488569
3	0.656294	0.276240
4	0.869922	0.321512
5	0.839483	0.363636
6	0.875545	0.371106
7	0.751834	0.392372
8	0.872668	0.409552
9	0.780505	0.412859
10	0.724707	0.395562

Table 4.2 Validity index of data C_2

c	TFL-SOC	TFL-HCM
2	0.615840	0.630156
3	0.817821	0.369271
4	0.500801	0.346411
5	0.859951	0.351662
6	0.647828	0.351353
7	0.689469	0.367606
8	0.656599	0.389425
9	0.734082	0.383173
10	0.755806	0.399439

Table 4.3 Validity index of data C_3

c	TFL-SOC	TFL-HCM
2	0.844930	0.662083
3	0.602924	0.647055
4	0.685708	0.481458
5	0.690660	0.393397
6	0.516887	0.301561
7	0.686904	0.440707
8	0.655841	0.335951
9	0.655931	0.366870
10	0.687098	0.379181

For all of these data sets, the optimal cluster numbers can be obtained correctly by the method of TFL-HCM and TFL-SOC. In Table 4.1, the validity index of data set C_1 is shown. The optimal cluster number, bold characters, becomes three clusters (c^*=3). In Table 4.2 and Table 4.3, the validity index of data sets C_2 and C_3 are shown. The correct results with TFL-HCM and TFL-SOC can be obtained.

Chapter 5
Atmospheric Laser Radar Clustering and Indexing System

5.1 Introduction

Recently, a number of laser radars have been developed in order to observe the behavior of the atmosphere. This system can measure the various conditions of ozone, vapor water, cloud, volcanic eruption and so on [1], [2]. Under ordinary circumstances, the atmospheric laser radar data are obtained as the 2-d image in the whole measurement period [7]-[9]. In other words, the obtained data become quite large in size and thus it becomes difficult to locate some important characteristics among the observed data quickly. As one of the solutions, an automatic indexing system is considered for data searching, sorting and selection.

Various techniques have been developed for the processing of a database with the results of data classification or data recognition. One is the artificial neural network which has been studied for several years, particularly in the field of speech/image classification and recognition [10]-[15]. The main characteristics of neural networks are generalization and learning, for any data. When we consider automatic data clustering, several critical initial conditions are usually required although almost all developed clustering networks work quite well with these conditions. It is not easy to obtain these conditions as a priori information.

Cluster validity indexing is an important measurement technique which is crucial for practical application for clustering. There have been some existing cluster validity criteria, which have been discussed in the previous chapter and also the cluster validity criterion based on a total fitting level method (TFL) has been proposed.

5.2 Basic Background of Lidar Data Characteristic

When the laser radar is used to probe or to study the behavior of the atmosphere, it has been called lidar (light detecting and ranging). There are several types of lidar used in this field [1], [2]. Each system has specific techniques and components depending on the purpose of measurements such as the measurements of ozone, vapor water and clouds, including aerosol measurements. Normally, aerosols are generated due to human activities and are located in the boundary layer. The laser radar may use only single wavelength or multi-wavelengths for complex systems. All systems use the various laser properties to interact with particle concentrations in the atmosphere based on the absorption, scattering, depolarization or Doppler frequency shift. The developed system described in this report applies the scattering properties to locate and image the aerosol distributions at around a 3 km height from the ground surface. The laser radar system is shown in Fig. 5.1 and the details of the system have been explained in the chapter 3.

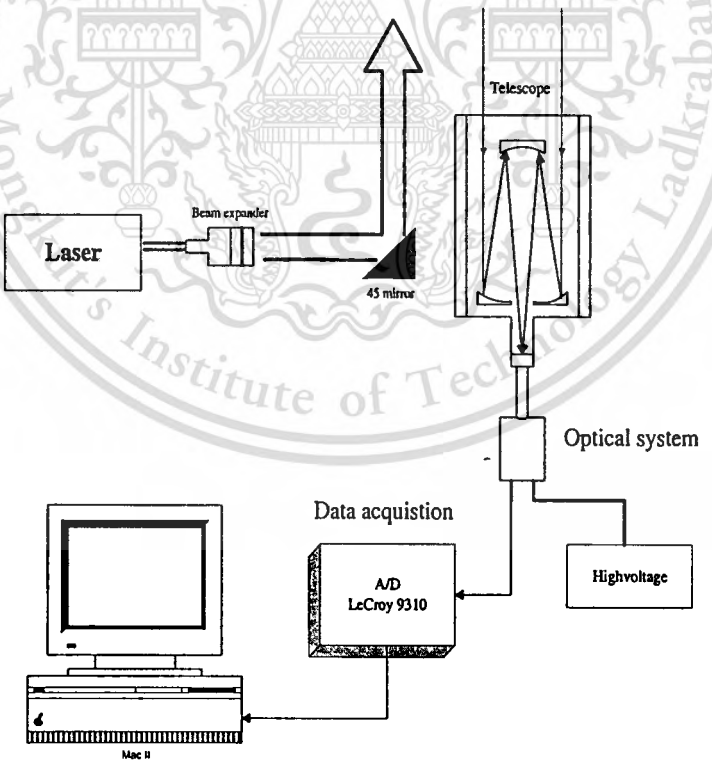
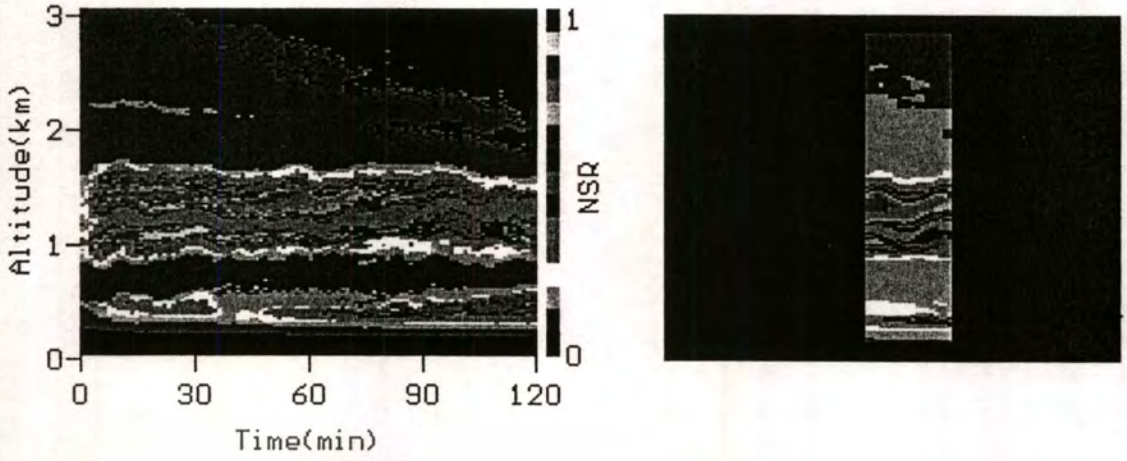


Figure 5.1 Block diagram of laser radar system

It is clearly recognized that the earth is covered by various mixed gasses which are divided into several layers with respect to the characteristics of the atmospheric temperature. The lowest atmospheric layer is called the troposphere in which many phenomena such as rain, snow, cloud, fog or aerosol occur. Moreover, the boundary layer is the region in the troposphere layer, i. e., only 3 km above the ground surface. Various pollutions that are generated by human activities affect this region. The laser has been used to explore the composition of mixing particles including the pollutions and aerosols in the atmosphere.

In this study, the ground-based laser radar has been used to locate the pollutants or aerosol distributions based on the properties of light scattering [2], [8]. If the data are continuously observed, the scattering ratio profile is obtained as shown in Fig. 5.2. In Fig. 5.2, the vertical axis denotes altitude, the horizontal axis denotes time and the colors indicate the normalized unit scattering ratio (NSR) at sixteen levels. These profiles show the located aerosols which vary with time and space [7], [9].

When the observation period is significantly long, the data records are increased and generate a very large size database. The characteristics of data are considered as the location of aerosols with thickness and altitude with respect to the ground surface. We require a method which is used to manage these data for clustering and indexing [22], [23]. The data are considered in the two-dimensional small section that is used for clustering. $f(a, b)$; a=1,2,...., 10 and b=1,2,...., 90. Each pixel is described with sixteen levels.



(a) Observed data (b) Small section

Figure 5.2 Characteristic of atmospheric laser radar data

5.3. Self-Organizing Clustering Network

Let us consider the set of k vectors in the s-dimensional feature space $X=\{y_1, y_2,..., y_k\} \subset R^s$ where the element of y_j is 900 given as f(a,b), i.e., y_j= [f(1, 1) f(1, 2)....f(10, 1) f(10, 2) f(10, 3).....f(10, 90)]. These vectors are assumed to be clustered into c groups. We define the family of cluster sets $\{A_i, i=1, 2,..., c\}$ for X where the following equations are satisfied:

$$\bigcup_{i=1}^{c} A_i = X$$

$$A_i \cap A_j = \phi \quad \text{all } i \neq j$$

$$\phi \subset A_i \subset X \quad \text{all } i$$

The number of A_i is clearly given by c, $2 \leq c \leq k-1$. Let us define c^* as the number of the optimal clusters that can be obtained from the cluster validity criteria.

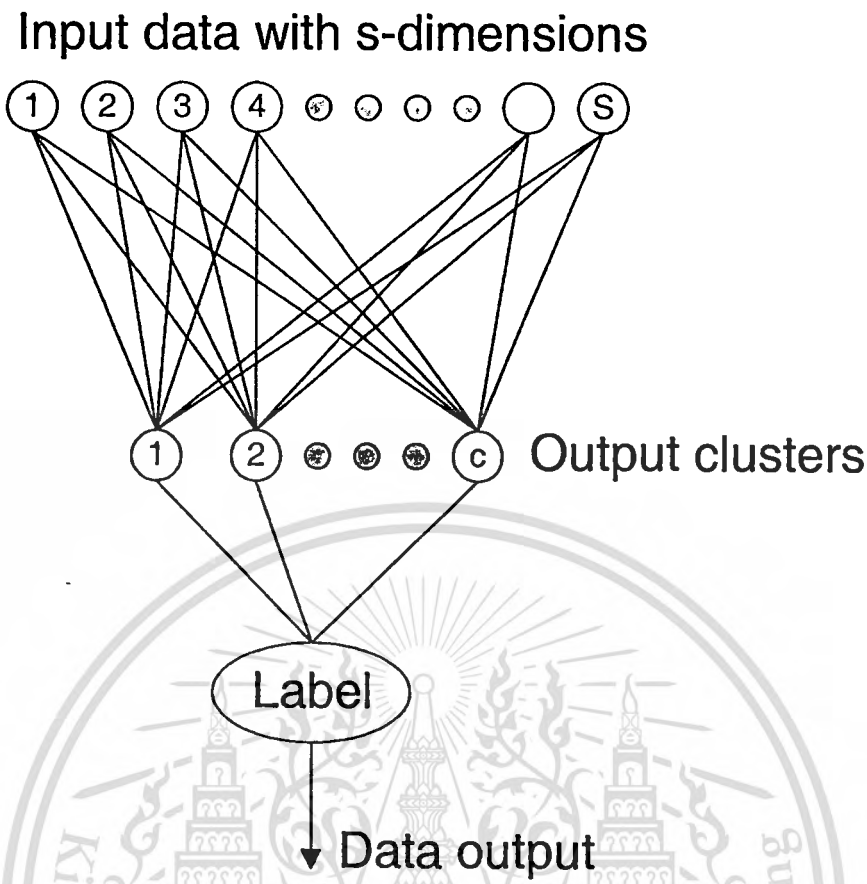


Figure 5.3 Self-organizing neural network

In Fig. 5.3, the basic construction of a self-organizing neural network is shown with the s-dimensional input data and c output clusters. The parameters of each cluster are given as random initial values and adapt with a self-organizing algorithm. The adaptive learning step of this algorithm is illustrated in Fig. 5.4. With this process, an input vector normally belongs to a selected cluster which has the largest similarity to the input vector. However, if the maximum similarity between the input vector and the information of clusters based on Euclidean distance is less than the appropriate threshold, i.e., T_s, a new cluster is generated.

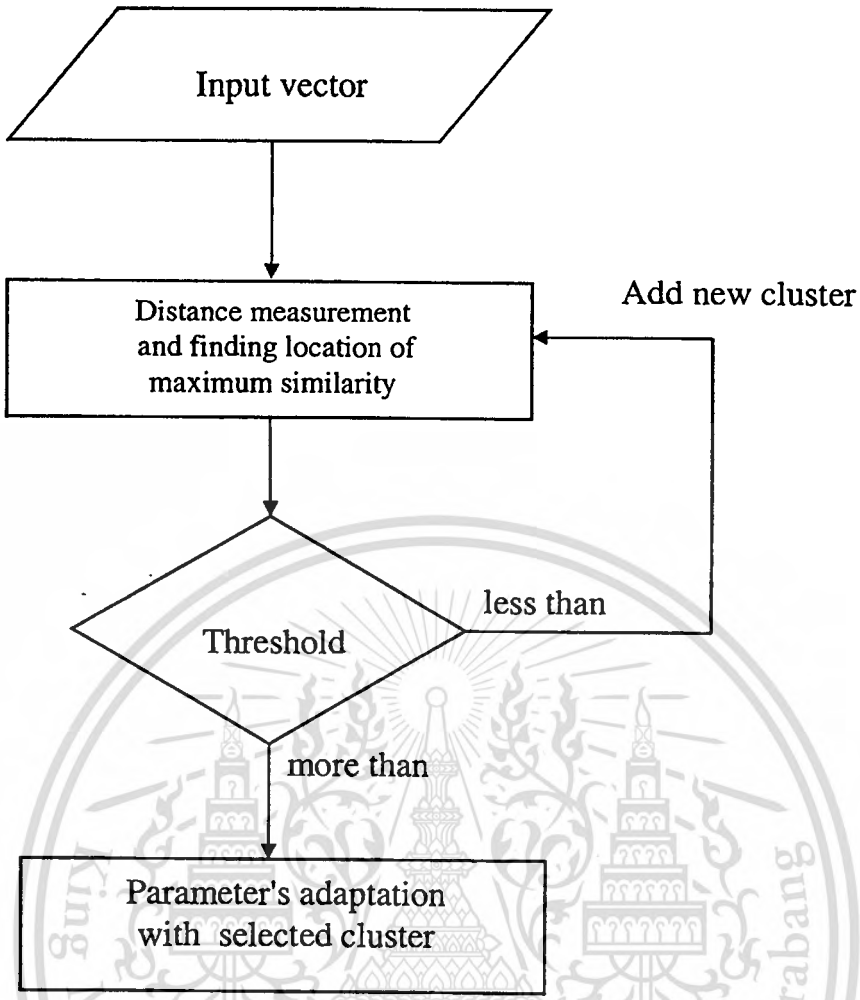


Figure 5.4 SOC learning

For an input vector, i.e., y_j (j= 1,2,...., k), each cluster determines whether this input belongs to the cluster or not. The input is determined as the member of the selected cluster which shows the maximum similarity. After the relative distances have been calculated in all clusters, the selection of the maximum similarity is executed with the index of selection $\delta(j)$ of $y(j)$; (j=1,2.....k). If $y(j) \varepsilon A_i$, $\delta(j)$ =1. Otherwise $\delta(j)$=0. The information of the cluster can be evaluated as

$$u_i(k) = \sum_{j=0}^{k} \delta_i(j) \qquad (5.1)$$

$$t_i(k) = \sum_{j=0}^{k} \Gamma^{u_i(k)-u_i(j)} \delta_i(j) \tag{5.2}$$

$$\mathbf{x}_i(k) = \frac{1}{t_i(k)} \sum_{j=0}^{k} \Gamma^{u_i(k)-u_i(j)} \delta_i(j) \mathbf{y}(j) \tag{5.3}$$

where Γ is a forgetting factor which has been set as $0 < \Gamma < 1$. Using this factor, every cluster considers the new data to be more important than the older data. This forgetting factor is useful for an adaptation mechanism and the elimination of random initial conditions. The variance $\mathbf{V}_i(k)$ of the cluster is calculated as

$$\mathbf{V}_i(k) = \frac{1}{t_i(k)} \sum_{j=0}^{k} \Gamma^{u_i(k)-u_i(j)} \delta_i(j) \cdot \mathbf{e}_i(j \mid k)\ \mathbf{e}_i(j \mid k)^T \tag{5.4}$$

where the error vector $\mathbf{e}_i(j|k)$ is defined as

$$\mathbf{e}_i(j|k) = \mathbf{y}(j) - \mathbf{x}_i(k) \tag{5.5}$$

and each element is assumed to be independent of others.

The measure of similarity is based on the standard euclidean distance in this process. In order to calculate the similarity values $d_i(\mathbf{y},k)(i=1,2,\ldots,c)$, each cluster has the information of a cluster pattern $\mathbf{x}_i(k)$, an error variance matrix $\mathbf{V}_i(k)$ and the value of member $t_i(k)$. The similarity between input data and cluster is finally defined as

$$d_i(\mathbf{y},k) = \exp[-(\alpha/2s)\mathbf{e}_i(k \mid k-1)^T \mathbf{V}_i^{-1} \cdot (k-1)\mathbf{e}_i(k \mid k-1)] \tag{5.6}$$

where α is a rectification factor [24], [25].

After the selected cluster is obtained with the maximum similarity, this value is compared with a threshold level. If it is less than the threshold level, a new cluster is generated. Otherwise, the information of the selected cluster with the maximum similarity is updated [24]-[26].

5.4 Cluster Validation

With the algorithms in the previous section, even when the mathematical forms for adaptations are obtained, the appropriate values of parameters are not given because of two initial conditions. We refer to the SOC learning in Fig. 5.4, where the number of clusters increases if the distance measurement of the maximum similarity is less than the threshold level. Moreover, since the forgetting factor in (5.2)-(5.4) directly affects the cluster information, the cluster number is changed with the forgetting value. Then, the result yields different clusters, $\{c_1, c_2, \ldots, c_n\}$, illustrated in Fig. 5.5.

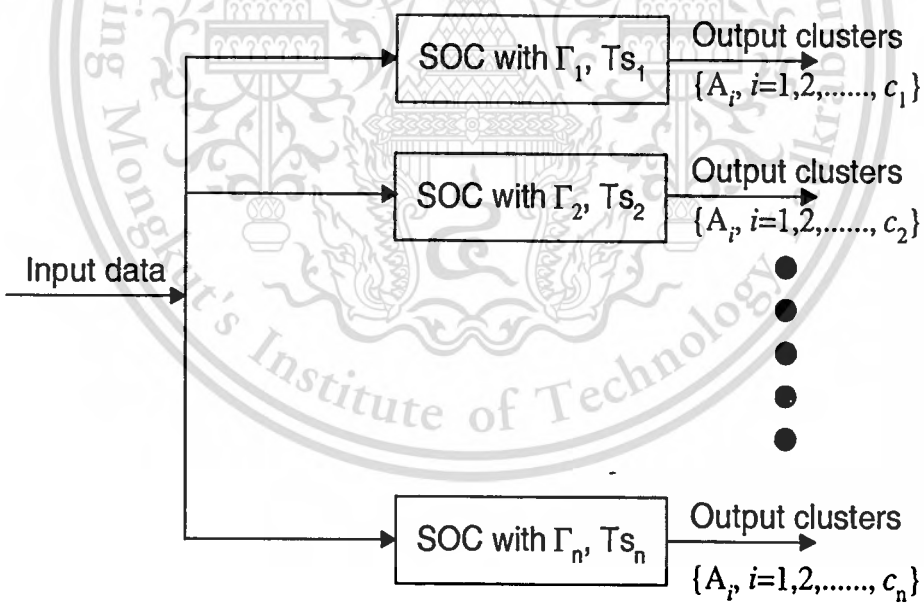


Figure 5.5 The cluster numbers of SOC learning

For this reason, an important question is how many clusters should there be. The cluster validity becomes a method that can be used to

determine the optimal cluster number. In the previous chapter, the total fitting level (TFL) cluster validity index has been used to determine the optimal cluster number based on the difference between the estimated normal probability distribution function P(k) and a data histogram h(k) of that cluster [21]. The fitting level at the i^{th} cluster is defined as

$$FL_i = \sum_{k=-\infty}^{\infty} |p_i(k) - h_i(k)| \tag{5.7}$$

The validity index function TFL(c) is the average of the fitting level at every cluster,

$$TFL(c) = \frac{1}{c}\sum_{i=1}^{c} FL_i \tag{5.8}$$

Figure 5.6 shows the characteristic between the total fitting level and clusters. The optimum number of clusters can be given at point c*.

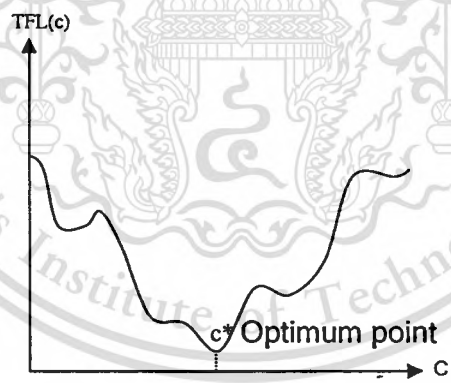


Figure 5.6 Validity index function TFL(c)

In experiments, it is not easy to determine the total fitting level because of several critical conditions. The parameters are determined by a simulated annealing method which can be described by the following steps:

Step 1

Given the input vectors. The forgetting factor Γ; $0<\Gamma<1$ and threshold level T_s; $0<T_s< 1$ are randomly given.

Step 2

SOC learning in Fig. 5.4 with (5.3), (5.4) and (5.6). Obtain the output cluster number.

Step 3

Compute the fitting level (FL) and total fitting level (TFL) using (5.7) and (5.8).

Step 4

Find the low value of the TFL for all random generations and then determine the narrower range of the forgetting factor and threshold level.

Step 5

At this point, the system has the narrower ranges of the parameters. Go to step 2 and step 3 for all of the new parameter ranges. The estimated parameters are obtained by the minimum value of the TFL.

5.5 Experimental results

In the algorithm, the clusters depend on the threshold and forgetting factor. For atmospheric laser radar data, Fig. 5.7 shows the result of clusters generated with T_s and Γ between 0.1-0.9. The maximum number of clusters, c=77, has been obtained at T_s= 0.9 and Γ = 0.1 and the minimum number of clusters, c=3, has been obtained at T_s= 0.1 and Γ = 0.9.

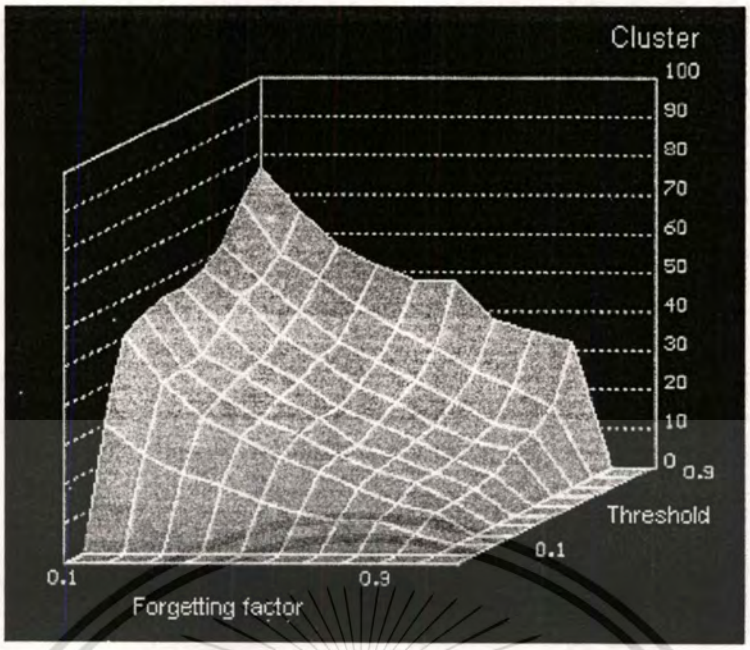


Figure. 5.7 Cluster generation with Γ, T_s

Unless the optimum point can be estimated, the arbitrary threshold and forgetting factor should be used for the calculation. Figure 5.8 shows the result of random generations versus the total fitting level. The circles at the i^{th} random generations in this figure mainly indicate the cluster number with the low total fitting level.

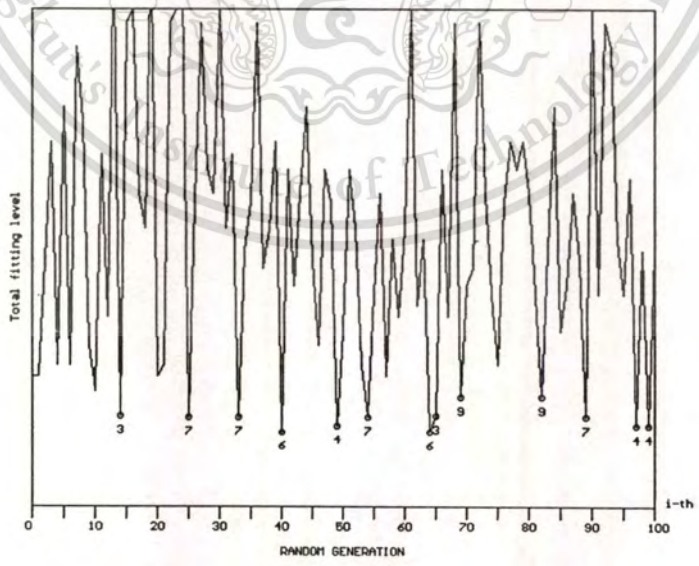


Figure 5.8 Random cluster generations

From the result of random generations in Fig. 5.8, the number of clusters is limited to between 3-9 clusters which correspond to the forgetting factor between 0.55-0.85 and the threshold value is less than 0.1. As for the threshold, the total fitting level is not affected when $T_s \leq 0.1$. According to this result, T_s is finally determined as 0.1. The result of iterative calculations is obtained and illustrated in Fig. 5.9. Finally, the optimum number of clusters is six, $c^*=6$, with $\Gamma = 0.68$ and $T_s = 0.1$.

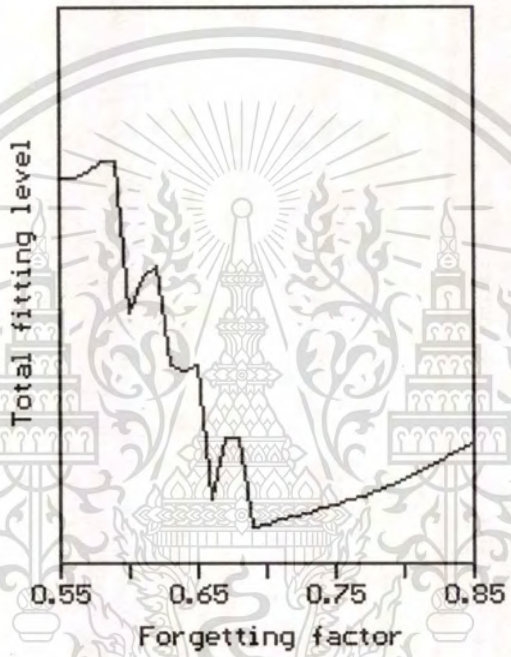


Figure 5.9 Optimum point

For the result of the given learning data, the parameters of each cluster have been adapted with appropriate conditions of the threshold and the forgetting factor given by TFL. The neural network system generates six clusters. The real data of atmospheric laser radar are illustrated in Fig. 5.10 as one example. The distances of (5.6) for all clusters are depicted in Fig. 5.11.

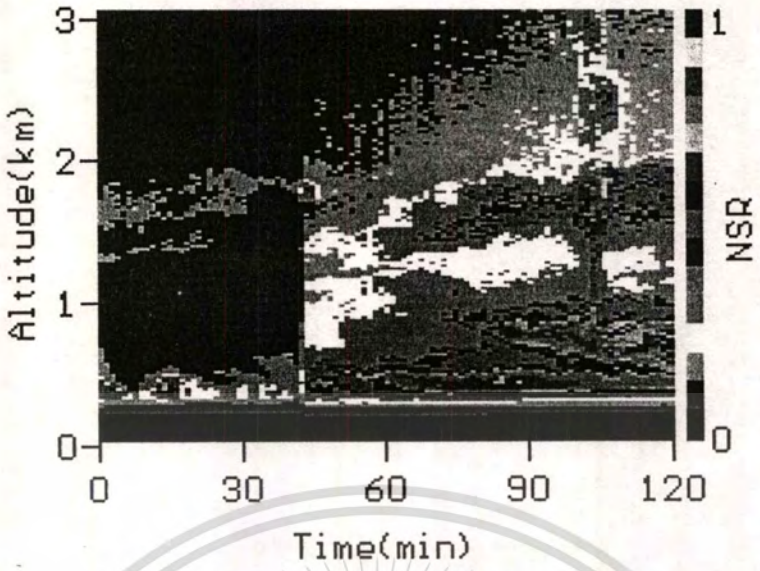
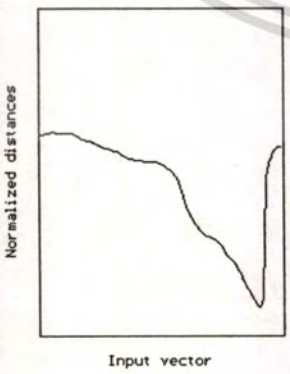


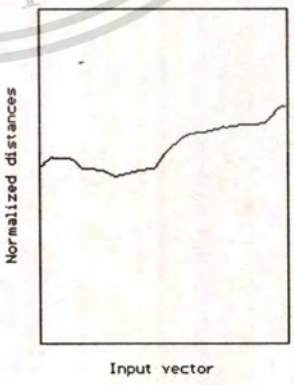
Figure 5.10 Input data for distance measurement

Cluster 1

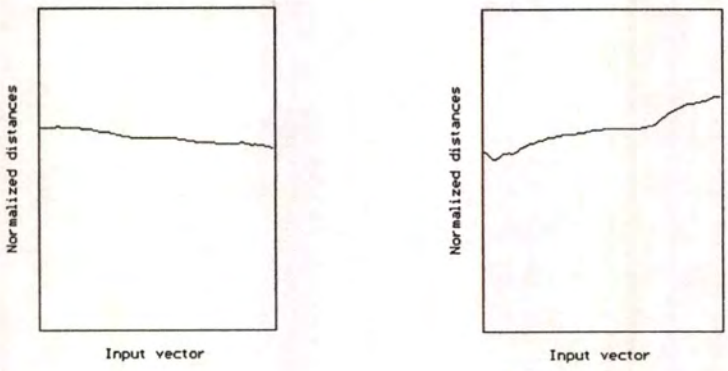
Cluster 2



Cluster 3



Cluster 4



Cluster 5 Cluster 6

Figure 5.11 Distance measurements of six clusters with input data

As another experiment, thirty data types are given as a new set. The input data are shown in Fig. 5.12 [22], [23]. The new set is automatically distinguished by using the already trained and self-organized clusters. It is shown in Table 5.1.

Table 5.1 Results of self-organizing clustering

Index	Input Vector
Group 1	No.1 No.6 No.15 No.24
Group 2	No.3 No.8 No.13 No.17 No.21 No.27 No.30
Group 3	No.12 No.20
Group 4	No.4 No.11 No.14 No.25
Group 5	No.2 No.5 No.9 No.18 No.22 No.26 No.28
Group 6	No.7 No.10 No.16 No.19 No.23 No.29

Figure 5.12 Input data set

Normally, the aerosols or the pollutants occur in the boundary layer and vary with time and space. In the clustering of Group 1 and Group 4, the locations of particles are appeared below 300 and 800 meters height, respectively. In the clustering of Group 2, the particles are located at around an 800 meters height and the particles layer has the thickness of approximately 400 meters. On comparing Group 3 and Group 6, it is seen that the particles appear in two layers separately but the upper layer has smaller thickness than the lower layer in Group 3. On the other hand, the lower layer has smaller thickness than the upper layer in Group 6. In the case of all data in Group 5, the pollutants content was very high, above

approximately a 2.5 km height. According to these indexing data, Group 5 is maximally polluted among all groups. The pollution level decreases for Groups 5, 6, 2, 4, 3 and 1. In terms of the pollution level, Groups 3 and 6 are highly similar. However their other characteristics are significantly different. For more detailed information, additional equipment is required.

Chapter 6

Conclusions

In this thesis, the first study and experiment of lidar system to obtain the aerosol profiles based on scattering ratio between Rayleigh backscattering and Mie backscattering have been done based on inversion method. However, to assume that the composition of particles in atmosphere consist only of gaseous molecules and aerosol. In the fact, many particles are in the boundary layer. For specific particle study in the atmosphere, there are several techniques, which need complex instruments. In the future, it is a need to develop the lidar system in other techniques.

This thesis has also considered data classification. The unsupervised clustering methods, Hard c mean and Self-organizing clustering, have been adopted in the process. One of the major problems encountered during the clustering of the given data is that the number of clusters is not known a priori. For this problem, an algorithm is required, that is a total fitting level method (TFL), for the cluster validity criterion. By TFL, the system measures the difference between a data histogram and the cluster probability distribution. The experimental results show that the TFL(c) index gives the minimum value at the optimum number.

Finally, the unsupervised algorithm for atmospheric laser radar data clustering and indexing system has been presented in chapter 5. The adopted algorithm is self-organizing clustering (SOC). During learning, the total fitting level (TFL) method is used for cluster validation. The validity index function TFL(c) gives the minimum value at the optimal clusters. By the TFL method, the system can learn by itself to find both an optimal number of clusters and the exact value of two unknown parameters. With the TFL method, the cluster optimization, $c^*=6$, has been obtained with $\Gamma = 0.68$ and $T_s = 0.1$. From the experiments of real data, this data processing can distinguish the whole thirty data patterns

into six indexing groups automatically. Moreover, its results are acceptable. When this method is used for long time observations, it is possible to obtain the new information such as frequency occurred of some data patterns, season changing or the characteristic of trend.

References

[1] Zuev E. and I.E. Naats I. E. Inverse Problems of Lidar Sensing of the Atmosphere, New York: Springer-Verlag, 1983.

[2] Measure R. M. Laser Remote Sensing: Fundamentals and Application, NewYork: Wiley, 1984.

[3] Fredriksson K. and Hertz H. M. "Evaluations of DIAL technique for studies on NO_2 using mobile lidar system." Appl. Optics, Vol. 23, No. 12, pp. 1403-1414, 1984.

[4] Browell E. V. "Airborne lidar measurements." The Review of Laser Engineering, Vol. 23, No. 2, pp. 49-61, 1995.

[5] Granatstein V. L., Rhinewine M. and Fitch A. H. "Remote sensing of gas concentrations in smokestack emission." Appl. Optics, Vol. 12, No. 7, pp. 1511-1515, 1972.

[6] Thomason L. W. "Observations of an anomalous aerosol extinction made by SAGE II following the eruption of Mt. Pinatubo." Geophys. Res. Lett., Vol. 19, No. 8, pp. 2179-2182, 1992.

[7] Lerkvaranyu S., Dejhan K., Cheevasuvit F. and Tooprakai P. "Lower tropospheric aerosol profile observation by ground-based lidar at KMITL, Thailand." Proc. 18th Asian Conference on Remote Sensing, pp. P-14-1 - P-14-5, Kuala Lumpur, Malaysia, October 1997.

[8] Lerkvaranyu S., Dejhan K. and Cheevasuvit F. "Echo signal analysis of ground-based laser radar in vertical axis." Proc. 1999 IEEE International Symposium on Intelligent Signal Processing and Communication Systems, pp. 625-627, Phuket, Thailand, December 1999.

[9] Lerkvaranyu S., Dejhan K., Cheevasuvit F., Itabe T., Mizutani K., Aoki T. and Yaisu M. "Atmospheric boundary layer observation by ground-based lidar at KMITL, Thailand

(13°N, 100°E)." Proc. SPIE Vol. 4153-31, pp. 226-233, Sendai, Japan, September 2000.

[10] Tou J. T. and Gonzalez R. C. **Pattern Recognition Principle**. Reading, MA: Addison-Wesley, 1974.

[11] Jain A. K. and Dubes R. C. **Algorithm for Clustering Data**. Englewood Cliffs: Prentice-Hall, 1988.

[12] Grossberg S. **Neural Network and Natural Intelligence**. London: The MIT Press, 1988.

[13] McClelland J. L. and Rumelhart D. E. **Parallel Distributed Processing**. London: The MIT Press, 1988.

[14] Kohonen T. **Self-Organizing Map**. Berlin: Springer-Verlag, 1995.

[15] Shiha S. K. and Karray F. "Classification of underground pipe scanned image using feature extraction and neural-fuzzy algorithm." IEEE Trans. Neural Network, Vol. 13, No. 2, pp. 393-401, 2002.

[16] Duda R. O. and Hart P. E. **Pattern Classification and Scene Analysis**. New York: Wiley, 1973.

[17] Pal N. R. and Bezdek J. C. "On cluster validity for the fuzzy c-means model." IEEE Trans. Fuzzy Syst., Vol. 3, No. 3, pp. 370-379, 1995.

[18] Xie X. L. and Beni G. "A validity measure for fuzzy clustering." IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 13, No. 8, pp. 841-847, 1991.

[19] Fukayama Y. and Sugeno M. "A new method of choosing the number of clusters for the fuzzy c-means method." Proc. 5th Fuzzy Syst. Symp., pp. 247-250, June 1989.

[20] Kim D. J., Park Y. W. and Park D. J. "A novel validity index for determination of optimal number of clusters." IEICE Trans. Inf. & Syst., Vol. E84-D, No. 2, pp. 281-285, 2001.

[21] Lerkvaranyu S., Miyanaga Y., Dejhan K. and Cheevasuvit F. "An optimum unsupervised clustering based on total fitting level method." Proc. 2003 IEEE International Symposium on

Intelligent Signal Processing and Communication Systems, Awaji Island, Japan, December 2003.

[22] Lerkvaranyu S., Miyanaga Y., Dejhan K. and Cheevasuvit F. "Atmospheric laser radar data classifier and indexing based on optimal data clustering techniques." Proc. 2nd Int. Syms. Communication and Information Technology, Chonburi, Thailand, pp. 190-193, October 2002.

[23] Lerkvaranyu S., Miyanaga Y., Dejhan K. and Cheevasuvit F. "Automatic indexing system for atmospheric laser radar data." Proc. 2002 IEEE Asian Pacific Conference on Communication and System, pp. 237-240, October 2002.

[24] Miyanaga Y., Hiroshige M. and Tochinai K. "Parallel and adaptive clustering method suitable for a VLSI system." Proc. IEEE SCAS'91, 91CH3006-4, pp. 356-359, June 1991.

[25] Miyanaga Y., Okumura S. and Tochinai K. "On generalization and adaptation of a self-organized clustering." IEICE Trans. A, Vol. J75-A, No. 7, pp. 1207-1215, 1992.

[26] Miyanaga Y. and Tochinai K. "Parallel VLSI architecture for multi-layer self-organizing cellular network." IEEE Trans. Electronics, Vol. 3E76-C, pp. 1174-1181, 1993.

Publications of the author

[1] S. Lerkvaranyu, Y. Miyanaga, K. Dejhan and F. Cheevasuvit, "Self-organizing clustering for atmospheric laser radar clustering and indexing system," Journal of Signal Processing, Vol. 7, No. 5, November 2003.

[2] S. Lerkvaranyu, Y. Miyanaga, K. Dejhan and F. Cheevasuvit: An optimum unsupervised clustering based on total fitting level method, Proc. 2003 IEEE International Symposium on Intelligent Signal Processing and Communication Systems, Awaji Island, Japan, December 2003.

[3] S. Lerkvaranyu, Y. Miyanaga, K. Dejhan and F. Cheevasuvit, "Improving QPSK demodulation in direct conversion receiver with RBF neural network," Proc. 2003 IEEE International Symposium on Intelligent Signal Processing and Communication Systems, Awaji Island, Japan, December 2003.

[4] S. Lerkvaranyu, Y. Miyanaga, K. Dejhan and F. Cheevasuvit, "Automatic indexing system for atmospheric laser radar data," Proc. 2002 IEEE Asian Pacific Conference on Communication and System, pp. 237-240, October 2002.

[5] S. Lerkvaranyu, Y. Miyanaga, K. Dejhan and F. Cheevasuvit, "Atmospheric laser radar data classifier and indexing based on optimal data clustering techniques," Proc. 2nd Int. Syms. Communication and Information Technology, Chonburi, Thailand, October 2002.

[6] S. Lerkvaranyu, K. Dejhan, F. Cheevasuvit, T. Itabe, K. Mizutani, T. Aoki and M. Yaisu: Aerosol profile enhancement with digital image processing, Proc. 2nd International Workshop on Radar and Lidar Remote Sensing Technology in Asia, pp. 32-37, Tokyo, Japan, October 2000.

[7] **S. Lerkvaranyu**, K. Dejhan, F. Cheevasuvit, T. Itabe, K. Mizutani, T. Aoki and M. Yaisu, "Atmospheric boundary layer observation by ground-based lidar at KMITL, Thailand (13°N, 100°E)," Proc. SPIE Vol. 4153-31, pp. 226-233, Sendai, Japan, September 2000.

[8] S. **Lerkvaranyu**, K. Dejhan, A. Mayanon, T. Itabe and K. Mizutani, "Measurement and signal processing of cirrus clouds under clear atmospheric condition," Proc. International Workshop on Radar and Lidar Remote Sensing Technology in Asia, pp. 37-47, Tokyo, Japan, February 2000.

[9] S. **Lerkvaranyu**, K. Dejhan and F. Cheevasuvit, "Echo signal analysis of ground-based laser radar in vertical axis," Proc. 1999 IEEE International Symposium on Intelligent Signal Processing and Communication Systems, pp. 625-627, Phuket, Thailand, December 1999.

[10] S. **Lerkvaranyu**, A. Mayanon, T. Itabe and K. Mizutani, "Lidar activity in THAILAND," Summary of Report Meeting KMITL-MPT/CRL on Radar and Lidar Remote Sensing, pp. 103-108, Tokyo, Japan, December 1998.

[11] S. **Lerkvaranyu**, K. Dejhan, F. Cheevasuvit, "Density and Temperature Profile in the Lower Atmospheric at KMITL, Thailand," Proc. 19th Asian Conference on Remote Sensing, pp. R-15-1 - P-15-6, Manila, Philippines, November 1998.

[12] S. **Lerkvaranyu**, K. Dejhan, F. Cheevasuvit, "Moving Average Method for Time Series Lida Data," Proc. 19th Asian Conference on Remote Sensing, pp. R-16-1 - P-15-5, Manila, Philippines, November 1998.

[13] S. **Lerkvaranyu**, K. Dejhan, F. Cheevasuvit and P. Tooprakai, "Lower tropospheric aerosol profile observation by ground-based lidar at KMITL, Thailand," Proc. 18th Asian Conference on Remote Sensing, pp. P-14-1 - P-14-5, Kuala Lumpur, Malaysia, October 1997.

[14] **S. Lerkvaranyu**, K. Dejhan, F. Cheevasuvit, P. Tooprakai, "Height and thickness of cirrus cloud estimation using Nd:YAG laser based on optical remote sensing techniques," Proc. 18th Asian Conferenceon Remote Sensing, pp.G-7-1-G-7-6, Kuala Lumpur, Malaysia, October 1997.

About the author

Somkiat Lerkvaranyu received his B. Sc degree in Physics from King Mongkut's Institute of Technology Thonburi (KMITT), Bangkok, Thailand, in 1993 and the M. Eng degree in Electrical Engineering from King Mongkut's Institute of Technology Ladkrabang (KMITL), Ladkrabang, Bangkok, Thailand, in 1996. He is currently a lecturer at Telecommunications Engineering Department, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang. His research interests are in the areas of adaptive signal processing, nonlinear signal processing and pattern recognition/classification.

Address:

Telecommunications Engineering Department, Faculty of Engineering
King Mongkut's Institute of Technology Ladkrabang (KMITL)
Chalongkrung Road, Ladkrabang District
Bangkok 10520, THAILAND
E-mail: somkiat@telecom.kmitl.ac.th, klsomkia@kmitl.ac.th