UNIVERSITI PUTRA MALAYSIA

PARAMETER ESTIMATION OF K-DISTRIBUTED CLUTTER BASED ON FUZZY INFERENCE AND GUSTAFSON-KESSEL CLUSTERING

ATEFEH DAVARI

PARAMETER ESTIMATION OF K-DISTRIBUTED CLUTTER BASED ON FUZZY INFERENCE AND GUSTAFSON-KESSEL CLUSTERING

By
ATEFEH DAVARI

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfillment of Requirements for the Degree of Master of Science

August 2008

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of
the requirement for the degree of Master of Science

PARAMETER ESTIMATION OF K-DISTRIBUTED CLUTTER BASED ON FUZZY INFERENCE AND GUSTAFSON-KESSEL CLUSTERING

By

ATEFEH DAVARI

August 2008

Chairman: Mohammad Hamiruce Marhaban, PhD

Faculty: Engineering

The detection performance of maritime radars is restricted by the unwanted sea echo or clutter which is the vector sum of scattering from the sea surface. The echo is noise-like and is expected from a set of randomly moving scatters. Although the number of these target-like data is small, they may cause false alarm in maritime radar and perturb the target detection. K-distribution is known as the best fitted probability density function for the radar sea clutter. The accurate and fast parameter estimation of K-distribution for small number of sea clutter radar data is crucial task to avoid irreparable disasters.

A novel approach to estimate the parameters of K-distribution based on fuzzy inference has been proposed in the thesis. Takagi-Sugeno Kang (TSK) model has been chosen since human knowledge is unavailable to be captured, whereas the sea clutter for specific parameter can be easily generated. GK- clustering has been used in order to identify the membership function of the antecedent parts. Least Square Method has been utilized to estimate the parameter of the K-distribution, which is represented in the consequent part of the fuzzy inference system. For a real-time

implementation of the proposed method, vectorized programming technique has been implemented. In comparison with the conventional methods, this technique has less computational complexity, needs lesser time to train and estimates faster than any existing methods. Since the method is clustering based, some kind of pre-knowledge (rough estimation) is naturally stored in the structure of the TSK-fuzzy system and Least Square provides a mechanism to fine tune the consequent parameters.

The novelty of the proposed method is the incorporation of the clustering (as a pre-estimator) with the estimation process. The resultant estimator then overcomes the bottleneck of the existing methods and is capable of handling even a small number of data.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia
sebagai memenuhi keperluan untuk ijazah Master Sains

PENGANGGARAN PARAMETER SELERAKAN TABURAN-K BERDASARKAN INFERENS KABUR DAN PENGELOMPOKAN GUSTAFSON-KESSEL

Oleh

ATEFEH DAVARI

December 2007

Pengerusi: Dr. Mohammad Hamiruce Marhaban, PhD

Fakulti: Kejuruteraan

Pencapaian pengesanan radar maritim terhad disebabkan gema atau selerakan laut yang tidak diingini terhasil daripada jumlah vektor penyelerakan daripada kawasan permukaan lautan. Gema seakan hingar ini dijangkakan dari set pergerakan selerakan yang rawak. Walaupun jumlah data sasaran ini kecil, ia mampu menyebabkan amaran salah dalam radar maritim dan mencemarkan pengenalpastian sasaran. Taburan-K dikenali sebagai fungsi ketumpatan kebarangkalian yang terbaik untuk selerakan laut radar. Ketepatan dan kepantasan anggaran parameter Taburan-K untuk data selerakan laut yang kecil adalah tugasan yang genting untuk mengelakkan bencana yang tidak dapat dipulihkan.

Suatu pendekatan baru untuk menganggar parameter Taburan-K berdasarkan inferens kabur telah diketengahkan di dalam tesis ini. Model Tagaki-Sugeno Kang (TSK) telah dipilih memandangkan pengetahuan manusia tidak dapat dikenalpasti, tetapi selerakan laut untuk parameter tertentu boleh dihasilkan dengan mudah. Pengelompokan-GK telah digunapakai untuk mengenalpasti fungsi keahlian bahagian anteseden. Kaedah kuasa dua terkecil telah digunakan untuk menganggar

parameter Taburan-K, yang diwakilkan di bahagian konsekuen dalam sistem inferens kabur.

Untuk pelaksaaan masa sebenar terhadap kaedah yang diperkenalkan, teknik pengaturcaraan bervektor dilaksanakan. Jika dibandingkan dengan kaedah konvensional, teknik ini mampu mengurangkan kerumitan pengiraan, memerlukan masa yang kurang untuk melatih sistem dan menganggar lebih laju berbanding kaedah sedia ada. Oleh kerana kaedah ini berdasarkan pengelompokan, pra-pengetahuan (anggaran kasar) secara semulajadi disimpan di dalam struktur sistem kabur-TSK dan kaedah kuasa dua terkecil menyediakan mekanisme untuk talaan halus parameter konsekuen.

Sesuatu yang baru dalam kaedah yang disarankan ialah penggabungan pengelompokan (sebagai pra-penganggar) dengan proses penganggaran. Penganggar terhasil telah mengatasi masalah yang dihadapi kaedah sedia ada dan mampu berfungsi walaupun dengan jumlah data yang sedikit.

ACKNOWLEDGMENTS

In the name of Allah, the Most gracious, the Most merciful.

First and foremost, I would like to express my gratitude to Allah, for giving guidance and ability to complete my thesis.

I would like to give respectfully all my sincere thanks to my supervisor Dr. Mohammad Hamiruce Marhaban. The thesis is indebted his support, encouragement and useful comments. Thanks and appreciation is extended to the members of the supervisory committee Dr. Samsul Bahari Mohd Noor.

My special thanks to the members of engineering faculty and GSO of Universiti Putra Malaysia for their helps.

My greatest thanks go to my father and my mother who always have encouraged and supported me in each state of my life. Thanks my good brother, parents in law and entire family. My final thanks are for my husband who is unique and the thesis could not been done without his attention and best helps and thanks my sons, the dearest persons in my life.

Peace and blessing upon Prophet Mohammad (s.t.w.).

I certify that an Examination Committee met on 12 August 2008 to conduct the final examination of Atefeh Davari on her master of science thesis entitled "Parameter Estimation of K-Distributed Clutter Based on Fuzzy Inference and Gustafson-Kessel Clustering" in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the candidate be awarded the relevant degree. Members of the Examination Committee are as follows:

Sudhanshu Shekhar Jamuar, PhD
Professor
Faculty of Engineering
University Putra Malaysia
(Chairman)

Syed Javid Iqbal, PhD
Lecturer
Faculty of Engineering
University Putra Malaysia
(Internal Examiner)

Zairil Azhar Zaludin, PhD
Lecturer
Faculty of Engineering
University Putra Malaysia
(Internal Examiner)

Mohd Alauddin Mohd Ali, PhD
Professor
Faculty of Engineering
University Kebangsaan Malaysia
(External Examiner)

HASANAH MOHD. GHAZALI, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date :

This thesis submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirement for the degree of Master.
The members of the Supervisory Committee are as follows:

Mohammad Hamiruce Marhaban, PhD
Lecturer
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Samsul Bahari b.Mohd.Noor , PhD
Lecturer
Faculty of Engineering
Universiti Putra Malaysia
(Member)

AINI IDERIS, PhD
Professor/ Dean
School and Graduate Studies
Universiti Putra Malaysia

Date: 13 November 2008

DECLARATION

I declare that the thesis is my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously, and is not concurrently, submitted for any other degree at Universiti Putra Malaysia or at any other institutions.

ATEFEH DAVARI

Date: 6 November 2008

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

ANN	Artificial Neural Networks
CFAR	Constant False Alarm Rate
ECF	Characteristic Function Method
EM	Expectation Maximization
FLOM	Fractional Lower Order Moment
GBK	Generalized Bessel K Function Density
GK	Gustafson-Kessel
KCFE	Kernel Characteristic Function Estimator
MF	Membership Function
ML	Maximum Likelihood
MLMOM	Maximum Likelihood and Method of Moment
MOM	Method of Moment
OLS	Ordinary Least Square
PDF	Probability Density Function
RV	Random Variable
STD	Standard Deviation
TMD	Test Data Matrix
TSK	Takagi-Sugeno Kang

CHAPTER 1

INTRODUCTION

1.1 Background Overview

The detection performance of a maritime radar is typically restricted by the undesired sea echo or clutter. Sea echo or clutter is the vector sum of scattering from the sea surface within illuminated area. The movement of the scatterers (waves, ripples, etc.) results changing in relative phases of their separate echoes and yielding change in the total echo. The echo is often inevitable noise-like phenomena in a collection of randomly moving scatterers [1]. As an example, Figure 1.1 shows the random pixilated forms, including some clutters at the bottom of radar screen while they are not the real targets.

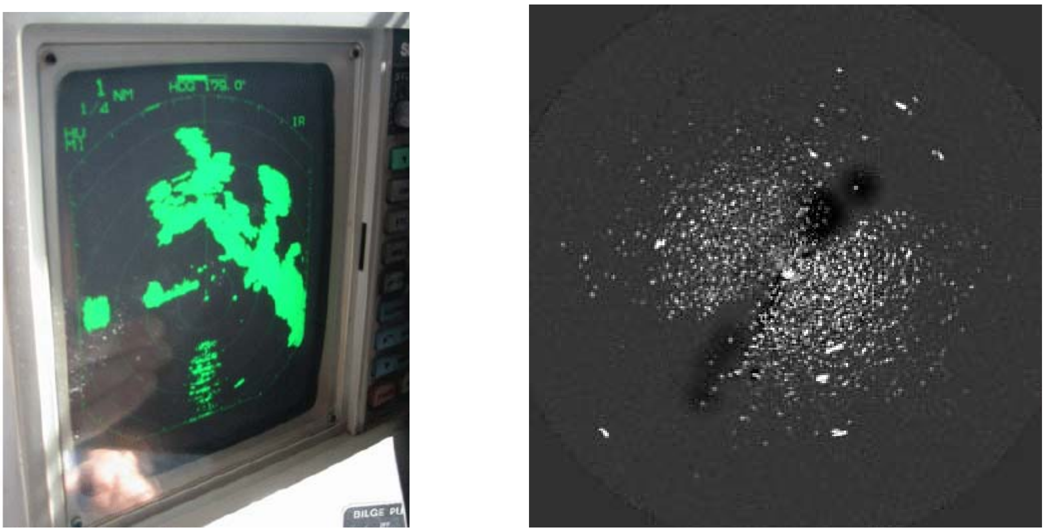

Figure 1.1. Random Scatterers Detected on Radar System

Simple models, like Gaussian or Rayleigh distribution can be applied effectively to model the sea clutter in low resolution radar signal. However, researchers have found that many non-homogeneous types of clutter, such as sea clutter, do not fit the Gaussian model well because of sea spike. These types of clutter lend themselves to a long tail in amplitude distribution; consequently, the conventional model does not perform well. Under these conditions, returns from the ocean can have target-like characteristics and may cause problem to distinguish from actual targets. In order to recognize between targets and clutter, and in some cases to prevent overloading radar operator or signal processor, the system must be capable to minimize constant false alarm rate (CFAR) from the clutter. Therefore more accurate model is needed and it has been found that other distributions namely Lognormal, Weibull, Contaminated-Normal, Log-Weibull and K-distribution can better fit the sea clutter. K-distribution model has been demonstrated to fit the sea clutter and was first time introduced by Jakeman and Pusey [2]. K-distribution has two parameters, namely, a and ν to be estimated. The scale parameter a can be calculated once and fixed during the estimation of shape parameter ν. Accurate estimation of K-distribution parameters is a crucial task in maritime targeting in which any error or delay in estimation may cause terrible consequents. Any inaccuracy in parameter estimation increases the probability of detection error so that the target is assumed as sea clutter or vice versa, mistakenly [3, 4]. It is also important to make the estimator capable of handling the small number of radar data since in practice the number of available radar data is small [5, 6].

1.2 Problem Statement

An accurate parameter estimation of K-distribution needs plenty of field data. However, in the real practice only small number of data is available for estimation [5, 6]. Therefore, the estimator needs to be capable of estimating accurately given small number of samples.

Figure 1.2 shows the scattering of estimated ν with high standard deviation in two traditional methods, FLOM and Raghavan's method [3].

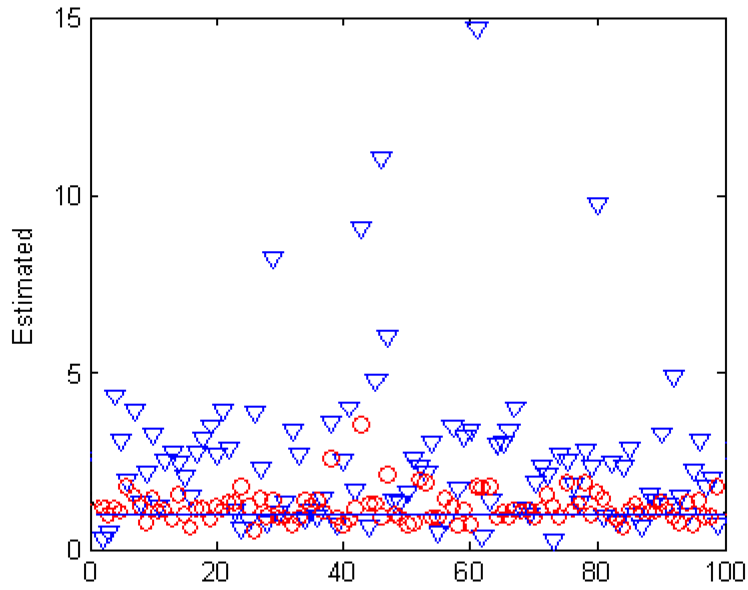

Figure 1.2. Independent trials of the estimates of ν for $N = 64$.
∇-FLOM $\bigcirc$-Raghavan's

In this situation, the probability of error is very high and the decision making algorithm will not have sufficient accuracy. Therefore, the system gives low performance of detection and every false alarm makes serious trouble. On the other hand, in the critical situations, the targeting would be performed as fast as possible. Then besides the accuracy, the researchers tried also to simplify the algorithm in order to reduce the program cycle time.

As a conclusion, the existing methods have looked into increasing the accuracy and simultaneously decreasing the computational complexity of the algorithm. However, the bottleneck of all of them is that they are not able to deal with a small number of data. Therefore an estimation algorithm is needed to be accurate and simultaneously fast (in comparison with the conventional methods) and furthermore capable of handling the estimation using a small number of data.

1.3 Aims and Objectives

The aim of this research is to design an accurate and fast estimator to estimate the parameter of the K-distribution, given a very small sample data. Combination of GK-clustering and TSK modeling has been chosen to take the advantage of clustering as pre-estimator and later fine-tuned by Least Square Regression. In order to achieve this aim the following objective are specified:

- **To develop a fuzzy function approximator using GK-clustering and TSK modeling (GK+TSK)**

  In this stage, first of all the range of universe of discourse of each fuzzy data is chosen to be the same as the range of data. The centers of clusters are then calculated by using GK-clustering. The center of clusters, now are assigned to be the basis vertices of fuzzy membership functions for the underlying fuzzy variables. In order to eliminate 'for' loops, shorten and fasten the fuzzification and fuzzy inference algorithms, vectorized programming is used.

- **To employ fuzzy function approximator (GK+TSK) as an estimator to estimate K-distribution parameter.**

  For this purpose, at first, statistics of data are chosen to be as fuzzy input variables. Then the fuzzy estimator is trained based on fuzzy inputs (statistics of K-distributed samples) and the fuzzy output (ν) which is supervisor. Then the trained fuzzy estimator is tested based on known samples and ν s.

- **To compare the proposed method with the previous methods**

  The thesis compares the proposed approach with the former methods in terms of performance, execution time and limitations.

  The worse limitation of previous methods is the inability to estimate based on small number of data. This problem has been overcome in the proposed method.

1.4 Outline of the Thesis

This thesis is made up of five chapters. The first chapter is introduction and includes overview background, problem statement, aims and objectives, outline of the thesis and scope of the work. This chapter ends with emphasizing of the contributions.

Chapter two starts with a critical review on the conventional estimation methods of K-distribution parameter. Then it briefs on the fuzzy logic, fuzzy modeling, TSK modeling and GK-clustering for determining the antecedent parts of fuzzy rules.

Chapter three is methodology that introduces a novel approach to estimate the K-distribution parameter using fuzzy TSK modeling and GK-clustering. This chapter demonstrates the complete algorithm of fuzzy estimator which includes generation of K-distribution data, defining the structure of TSK modeling, determining the center of membership functions using fuzzy GK-clustering, designing the membership functions and estimation of parameter of consequent parts of TSK rules. The chapter ends with elaboration of training and testing algorithm.

Chapter four, deals with the result and discussion. The first part of this chapter investigates the results of the proposed method in different numbers of sample data and in the next section the novel proposed method is compared with the conventional methods. Finally, the thesis is concluded in chapter five and also gives recommendations for the future works.

1.5 The Scope of Work

The scope of work of the thesis illustrated as follows:

- The thesis proposes a method to estimate the parameter of the K-distribution in order to model the sea clutter. It does not involve in the design of hardware or software of the radar system itself.
- In real practice, the scale parameter a can be calculated once the parameter ν has been estimated by the proposed method.
- The range of ν is considered between $0.1 - 1.5$ applicable in sea clutter radar data. The values of ν out of the mentioned range, is applicable in the ultrasonic medical applications and are not addressed through this thesis.

1.6 Contributions

The contributions of the thesis are highlighted as follows:

1. The thesis has proposed a novel estimator to estimate K-distribution based on fuzzy logic.
2. Unlike the former methods, the proposed method is able to estimate the parameter of K-distribution PDF, by using only a small numbers of data.

3. The method also has the learning property in which it can be trained in different predefined numbers of data to use in actual application.
4. The novel estimator has low bias and standard deviation in comparison with conventional methods.
5. The fuzzy estimator is fast and requires an execution time almost as small as the execution time of accurate previous methods because of two reasons;
    a. It is defined in compact matrix form and is developed based on vectorized programming approach.
    b. The method needs just one time training and afterward it can be employed for field data by using the matrix operations.
6. The proposed method has no singularity problem because in designing of membership functions in TSK modeling, shape of membership functions depend on the position of cluster centers rather than standard deviation of data. Therefore the universe of discourse is guaranteed to be complete and consistent.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction to K-distribution

In practice, the sea for high resolution radars can be modeled by two components: one of them is a spatially varying mean level y that results from a bunching of scatterers associated with the sea structure. The latter component is speckle which occurs due to multiple nature of the sea clutter and has fast fluctuation [1]. Based on these two components, the amplitude sea clutter distribution $p(x)$ is derived by averaging the speckle component over all possible values of the local mean level [4]:

$$p(x) = \int_0^\infty p(x \mid y) p(y)\, dy, \qquad 0 \le x \le \infty \tag{2.1}$$

where the speckle component $p(x \mid y)$ is Rayleigh distributed [4]

$$p(x \mid y) = \frac{\pi x}{2y^2} \exp\left(-\frac{\pi x^2}{4y^2}\right) \tag{2.2}$$

and $p(y)$ is the PDF of the clutter mean level, which has been found to be suitable fit to the Chi family of amplitude distributions[4].

$$p(y) = \frac{2b^{2\nu} y^{2\nu-1}}{\Gamma(\nu)} \exp\left(-b^2 y^2\right) \tag{2.3}$$

where b is the scale parameter and ν is shape parameter. Substituting Equations 2.2 and 2.3 into 2.1 yields the K-distribution [4]

$$p(x) = \frac{2}{a\Gamma(\nu)} \left(\frac{x}{2a}\right)^{\nu} \mathrm{K}_{\nu-1}\left(\frac{x}{a}\right) \qquad x > 0 \tag{2.4}$$

where $a = \sqrt{\pi}/2b$ is the scale parameter and ν is the shape parameter; $\Gamma(\cdot)$ is the gamma function and $K_\lambda(\cdot)$ is the modified Bessel function of order λ. For high resolution sea clutter, values of ν are generally observed in the region $0.1 \le \nu \le \infty$, where ν closed to 0.1 represents a very spiky clutter and $\nu = \infty$ represents thermal noise [3]. Figure 2.1 shows the amplitude PDF of K-distribution in different values of ν. The PDF of sea clutter represented by Equation (2.4) should be extracted from the radar data by estimation of parameters a and ν [3].

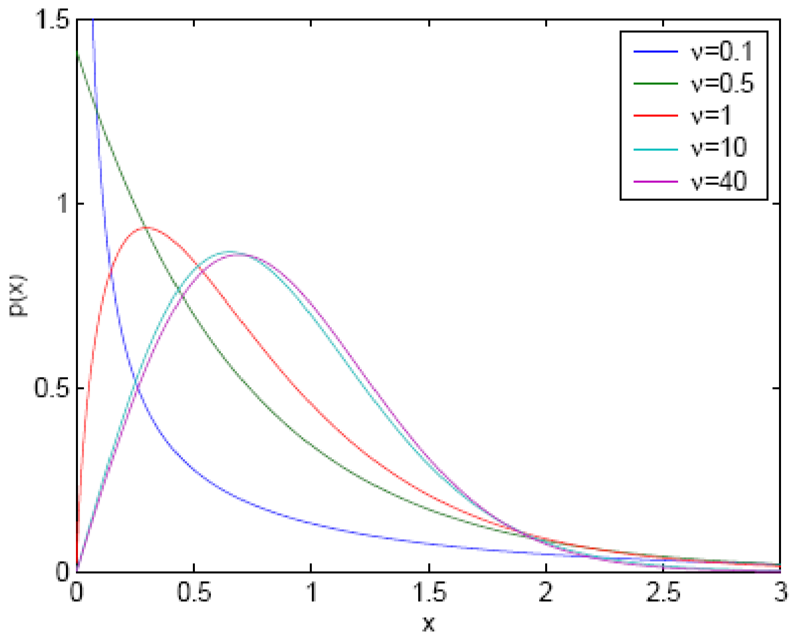

Figure 2.1. The amplitude PDF of K-distribution

2.2 Estimation of K-distribution Parameters

Maritime surveillance radars must detect very small targets among the background of sea clutter over a large area of oceans. Important application is in military industries; therefore the accurate detection of target is so important because a false alarm can make big trouble, so it entails essential assessment.