Natarajan Meghanathan
Jan Zizka (Eds)

Computer Science & Information Technology

8th International Conference on Artificial Intelligence, Soft Computing and Applications (AIAA 2018), November 24~25, 2018, Melbourne, Australia

AIRCC Publishing Corporation

Volume Editors

Natarajan Meghanathan,
Jackson State University, USA
E-mail: nmeghanathan@jsums.edu

Jan Zizka,
Mendel University in Brno, Czech Republic
E-mail: zizka.jan@gmail.com

ISSN: 2231 - 5403
ISBN: 978-1-921987-94-6
DOI : 10.5121/csit.2018.81601 - 10.5121/csit.2018.81607

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

Preface

The 8th International Conference on Artificial Intelligence, Soft Computing and Applications (AIAA 2018) was held in Melbourne, Australia during November 24~25, 2018. The 5th International Conference on Wireless and Mobile Network (WiMNeT 2018) and The 8th International Conference on Digital Image Processing and Pattern Recognition (DPPR 2018) was collocated with The 8th International Conference on Artificial Intelligence, Soft Computing and Applications (AIAA 2018). The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The AIAA-2018, WiMNeT-2018, DPPR-2018 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically. All these efforts undertaken by the Organizing and Technical Committees led to an exciting, rich and a high quality technical conference program, which featured high-impact presentations for all attendees to enjoy, appreciate and expand their expertise in the latest developments in computer network and communications research.

In closing, AIAA-2018, WiMNeT-2018, DPPR-2018 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the AIAA-2018, WiMNeT-2018, DPPR-2018.

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

Natarajan Meghanathan
Jan Zizka

Organization

General Chair

David C. Wyld	Southeastern Louisisna University, USA
Jan Zizka	Mendel University in Brno, Czech Republic

Program Committee Members

Abdolreza Hatamlou	Islamic Azad University, Iran
Adekunle	Ajasin University Akungba,Akoko, Ondo Nigeria
Aftab Alam	King Khalid University, Kingdom of Saudi Arabia
Alejandro Regalado Mendez	Universidad Del Mar, Mexico
Ali AL-zuky	Mustansiriyah University, Iraq
Ali Asghar Rahmani Hosseinabadi	Islamic Azad University Amol,Iran
Amit Choudhary	Maharaja Surajmal Institute, India
Ankit Chaudhary	Truman State University, USA
Ankur Singh Bist	Kiet Group of Institutions, India
Antonia Plerou	Ionian University, Greece
Asadollah Shahbahrami	University of Guilan, Iran
Atanu Nag	Modern Institute of Engineering & Technology, India
Avadhani P.S	Andhra University, India
Ayad Salhieh	Australian College of Kuwait, Kuwait
Benjamin Bedregal	Federal University of Rio Grande Do Norte, Brazil
Chidiebere Ugwu	University of Port Harcourt Choba, Nigeria
Chien-Cheng Yu	Hsiuping University of Science and Technology, Taiwan
Dongping Tian	Baoji University of Arts and Sciences, China
Doreswamy	Mangalore University, India
Emilio Jimenez Macias	University of La Rioja, Spain
Faiz ul haque Zeya	Bahria University, Islamabad, Pakistan
Farida Bouarab-Dahmani	Mouloud Mammeri University of Tizi-Ouzou, Algeria
Farshchi	Tehran University, Iran
Farzin Piltan	University of Ulsan, Korea
Fernando Bobillo	University of Zaragoza, Spain
Fernando Zacarias Flores	Universidad Autonoma de Puebla, Mexico
Grigorios N.Beligiannis	University of Patras, Greece
Hamid Ali Abed AL-Asadi	Basra University, Iraq
Hasnaoui Salem	Al-Manar University, Tunisia
Helmy	Universiti Tun Hussein Onn Malaysia, Malaysia
Hemalatha Nambisan	St Aloysius college, India
Hossein Jadidoleslamy	University of Zabol, Zabol, Iran
Hossein Khademolhosseini	Islamic Azad University, Iran
Houcine Hassan	Univeridad Politecnica de Valencia, Spain
Isa Maleki	Islamic Azad University, Iran
Islam Atef	Alexandria University, Egypt
Israa Shaker Tawfic	Ministry of Scientific and Technology, Iraq

Technically Sponsored by

Computer Science & Information Technology Community (CSITC)

Networks & Communications Community (NCC)

Soft Computing Community (SCC)

Organized By

Academy & Industry Research Collaboration Center (AIRCC)

TABLE OF CONTENTS

8th International Conference on Artificial Intelligence, Soft Computing and Applications (AIAA 2018)

An Intelligent Business Inventory Management Application Using Artificial Intelligence and Voice Recognition ……..... 01 - 07
Zehao Li, Yu Sun and Fangyan Zhang

Improved Block Stagewise Regularized Orthogonal Matching Pursuit Image Reconstruction Method ……..... 09 - 21
Xiong-yong Zhu, Shun-dao Xie, Guo-ming Chen, Liang Xue, Wen-fang Wu and Hong-zhou Tan

Ensemble Learning Based Voting Model for Dynamic Profile Classification and Project Allotment ……..... 77 - 88
Suhas Tangadle Gopalakrishna and Vijayaraghavan Varadharajan

5th International Conference on Wireless and Mobile Network (WiMNeT 2018)

Time-Domain Signal Management for OFDM Signals……..... 23 - 35
Takuya Kazama, Kazuki Miyazawa, and Masahiro Muraguchi

IoT-Based Home Appliance System (Smart Fan) ……..... 37 - 46
Mehran Ektesabi, Saman A. Gorji, Amir Moradi, Suchart Yammen, V. Mahesh K. Reddy and Sureerat Tang

A PAPR Reduction Technique in OFDM Systems with a Large Number of Subcarriers ……..... 47 - 61
Yasuhiro Shimazu, Yushi Shirato and Masahiro Muraguchi

8th International Conference on Digital Image Processing and Pattern Recognition (DPPR 2018)

Object Localization and Activities Identification Using Attribute Details in Smart Meeting Rooms ……..... 63 - 76
Dian Andriana, Ary Setijadi Prihatmanto, Egi Muhammad Idris Hidayat and Carmadi Machbub

AN INTELLIGENT BUSINESS INVENTORY MANAGEMENT APPLICATION USING ARTIFICIAL INTELLIGENCE AND VOICE RECOGNITION

Zehao Li[1], Yu Sun[2], Fangyan Zhang[3]

[1]University High School, Irvine, CA 92612
[2]California State Polytechnic University, Pomona, CA 91768
[3]Mississippi State University, Mississippi State, MS 39762

ABSTRACT

Virtual Enterprise, a class that simulates real-world business world, is designed for high school students to improve their experience in buying and selling. In Virtual Enterprise, various virtual products are created by classes. Each class is trying to sell products to students from different class in exhibitions. However, in trading, thousands of different types of handwritings made the salesperson in exhibitions are too difficult to be organized by administrator. To solve this problem, this paper develops an application, called Easy Exhibition, using Artificial Intelligence technology, which uses voice recognition technology to automatically complete the sales order by voice instead of by handwriting. The experiments show that AI-assisted solution improves both accuracy and efficiency in transaction processing.

KEYWORDS

Virtual Enterprise, Artificial Intelligence, Voice Recognition

1. INTRODUCTION

There is a class for high school students to attend called Virtual Enterprise [7] [13], which is a class for students to have a basic feeling on the business world especially the exhibition and buy and sell experiences. In this class, the whole class together creates a business company, and each person have an assigned role such as CEO, COO, CFO, and department heads. The company togethers creates the virtual product to sell and try to sell them to other students from different class of virtual enterprise from different school in exhibitions, which is also called trade shows [6]. However, the high school students do not have sufficient experiences to perform efficiently in those exhibitions, and therefore students suffer so much when sales are signing orders with customers on paper, that thousands of different types of handwritings made the salesperson have a really hard time to sort things out and send confirmation emails or sales receipts. Especially as a vice president of sales, that manages all the sales processes. Basically, my daily work in that class is recognize the contact information that customer wrote down on the sales receipt and try to send the email based on their terrible and rushing handwriting. Reading one is fine, but 100 every day is a pain in the neck.

A real company could also face this kind of problem on their real-world exhibition. When their customer makes purchases, it is almost impossible for sales person to record all kind of product

Natarajan Meghanathan et al. (Eds) : AIAA, WiMNeT, DPPR - 2018
pp. 01–07, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.81601

with every detail. And once the booth gets popular, the sales person will easily forget about which customer they are helping and what did they order one minutes ago.

Manually writing and typing the sales order is tedious, error-prone and time-consuming. With all kind of different handwritings that were wrote by people with big diversity in a rushing period, it turns out the writing are not recognizable at all. People's handwriting in a rushing period on an unstable platform makes the writing very messy, to the fact that nobody else can read what's written on the paper. It is so hard afterwards for salesperson to recognize the letter wrote on the paper, and the handwriting creates such a similar between letter that looks alike. Also, writing sales order by hand on paper takes so much time for both customers and salesperson. The customer need to fill in all kind of information such as name, phone number, email, and school by hand, and salesperson need to write down all product that customer purchased and calculate the price manually by a calculator on their phone. The salesperson cannot reach the customer, and the order they make are not valid anymore.

Now in the world that the Artificial Intelligence are growing rapidly, we can easily use the AI and voice recognition [8] [9] [10] technology to automatically complete the sales order by voice instead of the old inefficient way of complete them by hand. This way, it will make the salesperson complete the order faster and have a valid contact information to customers, and buyers can waste less time finishing the paperwork with salesperson once the purchase agreement has met.

2. MOTIVATION

Here's two examples (Figure 1 and Figure 2) of handwriting on the sales receipt that is hardly to recognize for salesperson. The letter is all scramble together, and some letter are uncertain whether it is a r or n, a 4 or 9, and whether there is a dot or not on the email section. Those elements are combined to create a contact information that is not readable. With this kind of contact information on the paper, the salesperson suffers so much after the exhibition when they are trying to contact consumers with order details. Manually writing and typing the sales order is tedious, error-prone and time-consuming. That creates a huge disadvantage for virtual enterprise company to manually fill out the sales receipt.

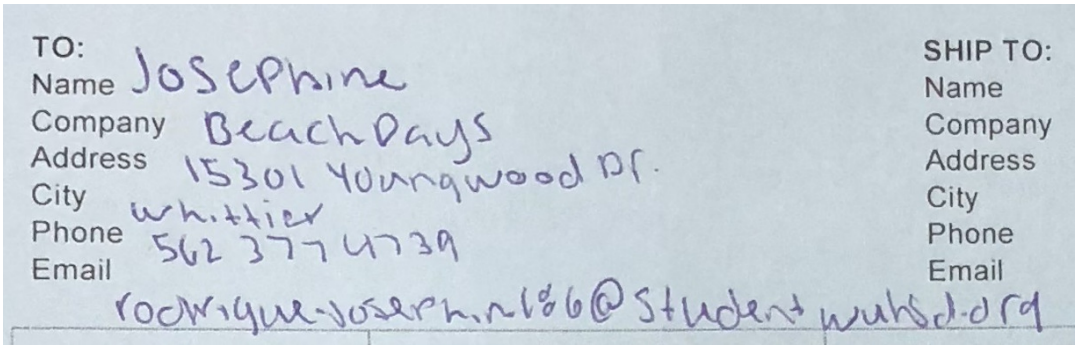


Figure 1: transaction written in handwriting

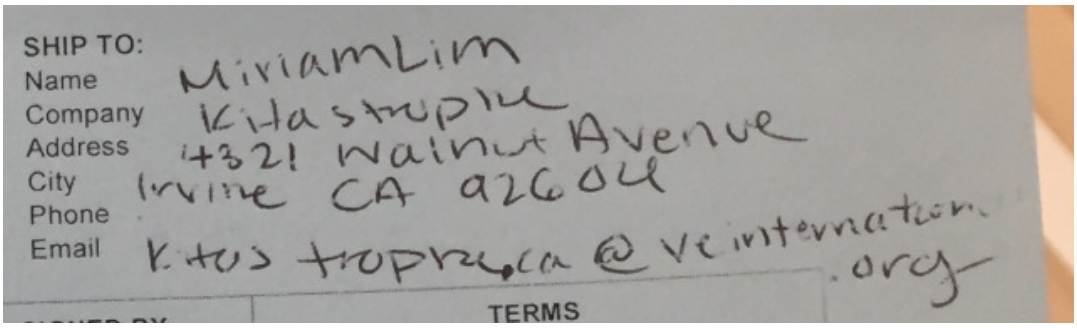


Figure 2: transaction written in handwriting

Different person in from different origin have different accent when they speak, and in this divergent nation all kind of accent exists. This creates a difficulty for voice recognition system to perform accurately on what the consumer says. Also, in the place where most of the voice recognition take place, there are a lot of different kind of unnecessary noise and might experience interference with the salesperson next to you. And, different category of information requires different way of voice input. For example, the company and school are one to five different word that customer and easily say the word, whereas the email is a series of random letter or number that requires to spell out every letter in order. The switch between saying the word and spell the letter and number make the voice recognition part hard to complete.

The user of the application is different companies, therefore there are different kind of products with different prices. Also, companies might locate in different states where the sales tax percentage is different. To set up a formula for calculating the price is confusing. And before this app, all the calculation is manually done by hand and a calculator on the phone, which is time consuming and error-prone.

The user needs to have the application that is going to help the business the best in term of time and effort. Some information on the sales order for company is just a simple number, that using voice recognition technique takes more time than handwrite a line on the paper to represent one.

3. SOLUTION

3.1. Overview of the Solution

The overview of the system is presented in Figure 3. The application created called Easy Exhibition, which uses voice recognition technology to substitute the customer information fill by hand. Also, the Easy Exhibition application can help the salesperson by automatically calculates the price and sends confirmation email to customers once the order being confirmed. The sales person only need to enter the quantity of product being sold and the application will help the salesperson do the rest of the work.

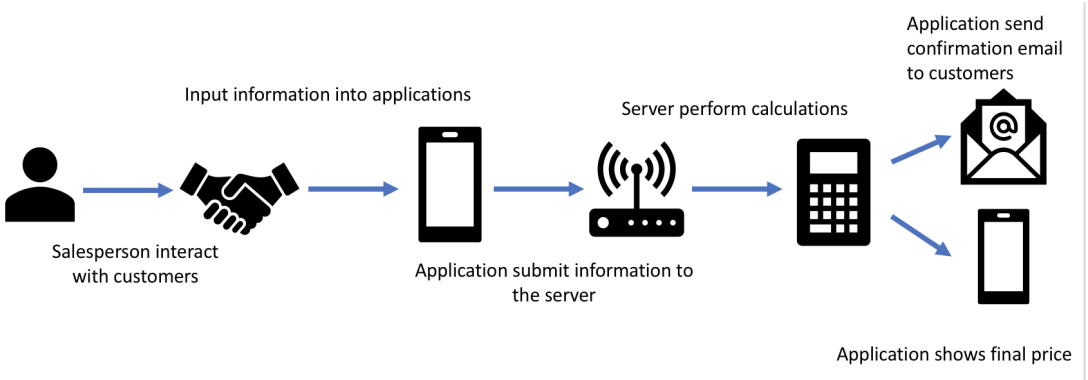


Figure 3: Architecture of Easy Exhibition

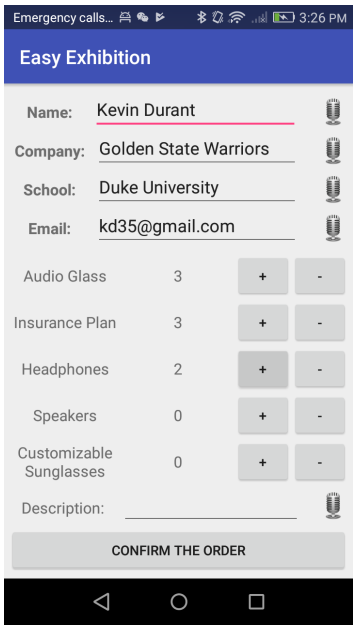


Figure 3: Screen shot of Easy Exhibition application

3.2. Voice Recognition

In the world of business, all information need to be completed and accurate. The voice recognition process in the app requires a high accuracy for the whole application function. In the app easy exhibition, we decided to use the voice recognition program established by Google. Based on my research, Google's voice recognition is one of the most popular and accurate voice input system.

3.3. Calculation System

On a trade show, not all the information is capable of using voice recognition system to operate. For example, the quantity of the product ordered is less efficiency to use voice recognition technology than a simple +/- button on the screen. Also, on the backend of the program, we use python calculation system [10] [12] to setup the equation to calculate the different kind of total price and taxation prices.

3.4. Interface

Not all the users for the business app are technical users, so the app (Figure 4) are created based on a friendly user interface. For the information that take a significant amount of time to write and hard to recognize on writing, the voice input system turns in to play. But for those section that only requires a number, an + or - button was used instead of voice recognition system to minimize and time taking to complete an order. This way, the final version of the sale receipt is clean and straightforward.

4. EVALUATION

To evaluate the performance of our system, we designed several experiments to test the performance of the system. Experiments are used to test accuracy, efficiency, and user experience.

A. The Accuracy of the Voice Input

To test the accuracy of the system, we tried to input 200 email address/names/school with 5 users. As shown in Figure 1, the average accuracy is 97.4%. Particularly, the accuracy of email recognition using voice is 97.3%, which is better than typing the email addresses manually.

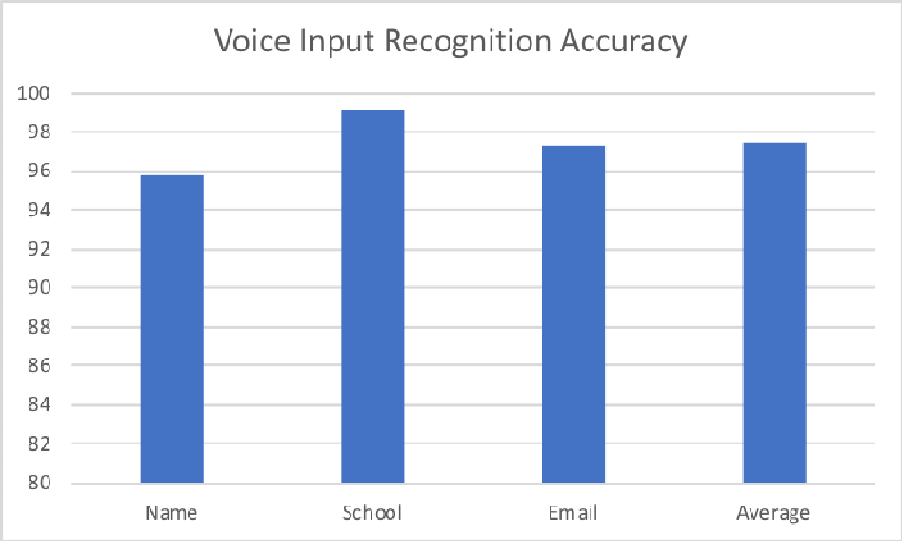
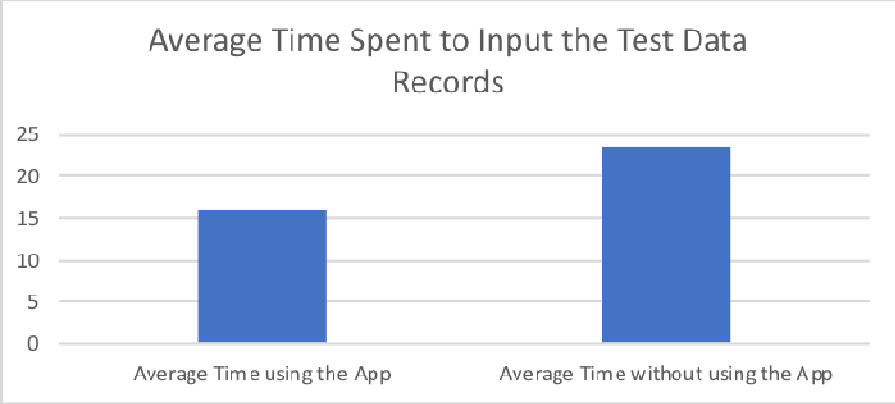


Figure 4. The Accuracy of the Voice Input Recognition Accuracy

B. The Improvement of the Efficiency

In this experiment, we divide the 6 users into 2 groups – one group use the app for data input, while the other manually type the same amount of information to an excel file. It shows that the average of time the App group is 32% faster, as shown in Figure 5.



C. The User Experience Feedback

We also created a questionnaire about the app and the user experience. 10 users participated provided the following feedback. As shown in Table 1, the overall experience from the users are very satisfying.

Question	**Average Rating (1 is the lowest and 5 is the highest)**
Does the app help you to input the information in a faster way?	4.5
Is the email voice input better than typing?	4.8
How accurate the information collected using the app is?	4.4
Is the app user-friendly?	4.8
Would you like to use the app to collect the customer information?	5

Table 1. The User Questionnaire Feedback

5. RELATED WORK

Similar work also applied machine learning to voice recognition. Muda, Begam [1], introduced one technique for voice recognition, which uses Mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) technique to improve voice recognition efficiency. Lancker [2] investigated recognition of familiar voices and presented 45 famous voice stimuli to 94 subjects, forward and backward. Their results showed that each familiar voice is best viewed as a relatively unique pattern. Human voice is done by adjusting different voices parameters. Until now, voice recognition is applied in various domains [3] [4] [5], which brings us great convenience in daily life.

6. CONCLUSION AND FUTURE WORK

In this project, we proposed an intelligent approach to address the problem of keeping record of transactions in Virtual Enterprise. This is realized by using voice recognition technology to automatically complete the sales order supported by machine learning. A mobile app has been developed to allow users to input orders through voice instead of handwriting.

To evaluate the performance of the application, we did several experiments for testing in accuracy, efficiency, and users' feedback. Experiments show that our accuracy can reach to high level in regular mode.

As for the future work, we will investigate other machine learning algorithms to keep improving the accuracy. We also would like to explore the possibility of applying deep learning in this problem domain.

In addition, one limitation related with the app is that its performance is dependent on google voice. We will look for better voice recognition backend to support our system. One feature we plan to add in the next version of the app is to upgrade interface and performance.

REFERENCES

[1] Muda, Lindasalwa, Mumtaj Begam, and Irraivan Elamvazuthi. "Voice recognition algorithms using mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques." arXiv preprint arXiv:1003.4083 (2010).

[2] Van Lancker, Diana, Jody Kreiman, and Karen Emmorey. "Familiar voice recognition: patterns and." Journal of phonetics13 (1985): 39-52.

[3] Miller, Stephen S. "Switching the modes of operation for voice-recognition applications." U.S. Patent 7,035,805, issued April 25, 2006.

[4] Lengen, Nicholas D. "Command insertion system and method for voice recognition applications." U.S. Patent 7,260,529, issued August 21, 2007.

[5] Miyazawa, Yasunaga, Mitsuhiro Inazumi, Hiroshi Hasegawa, and Isao Edatsune. "Interactive voice recognition method and apparatus using affirmative/negative content discrimination." U.S. Patent 5,899,972, issued May 4, 1999.

[6] Kijewski, Valerie, Eunsang Yoon, and Gary Young. "How exhibitors select trade shows." Industrial marketing management 22, no. 4 (1993): 287-298.

[7] Martinez, Ma Ta, Philippe Fouletier, Ka Ha Park, and Joel Favrel. "Virtual enterprise–organisation, evolution and control." International journal of production economics 74, no. 1-3 (2001): 225-238.

[8] Miller, Stephen S. "Switching the modes of operation for voice-recognition applications." U.S. Patent 7,035,805, issued April 25, 2006.

[9] Everhart, Charles A., Diane V. Gourd, Rajeev M. Joshi, and John Kefalos. "Advanced voice recognition phone interface for in-vehicle speech recognition applications." U.S. Patent 6,845,251, issued January 18, 2005.

[10] Zelle, John M. Python programming: an introduction to computer science. Franklin, Beedle & Associates, Inc., 2004.

[11] Shimada, Keiko. "Voice recognition dialing unit." U.S. Patent 5,222,121, issued June 22, 1993.

[12] Sanner, Michel F. "Python: a programming language for software integration and development." J Mol Graph Model 17, no. 1 (1999): 57-61.

[13] Hardwick, Martin, and Richard Bolton. "The industrial virtual enterprise." Communications of the ACM 40, no. 9 (1997): 59-60.

INTENTIONAL BLANK

IMPROVED BLOCK STAGEWISE REGULARIZED ORTHOGONAL MATCHING PURSUIT IMAGE RECONSTRUCTION METHOD

Xiong-yong Zhu[1], Shun-dao Xie[2,3], Guo-ming Chen[1], Liang Xue[1], Wen-fang Wu[2], Hong-zhou Tan[2,3]*

[1]Department of Computer Science,
Guangdong University of Education Guangzhou, China
[2]School of Electronic and Information Engineering
Sun Yat-Sen University Guangzhou, China
[3]SYSU-CMU Shunde International Joint Research Institute, Foshan, China

ABSTRACT

Traditional methods to signal acquisition need to collect large amounts of redundant data, and then compress the data to extract useful information, which is inefficient and requires large amount of storage resources. Compressed sensing (CS) can avoid sampling the redundant data; it obtains the discrete signals at the sampling rate that is lower than the Nyquist sampling rate, and reconstructs the original signal with high probability. Based on CS, Block Stagewise Regularized Orthogonal Matching Pursuit (StROMP) is proposed in this paper to reconstruct images. Simulation results show that the proposed algorithm can effectively reduce the required storage storages and computational complexity, which improves the quality of reconstructed images in the premise of ensuring a shorter reconstruction time.

KEYWORDS

Compressive Sensing; Matching Pursuit; Image Reconstruction

1. INTRODUCTION

The traditional method to signal acquisition is based on the theorem of Nyquist sampling. We usually proceed by all the original data and then compress it to get the actionable message. However, most of the data that we acquire can be thrown away with almost no perceptual loss. The classical way wastes a great deal of sampling resources, and requires more resources for data storage and processing. Then compressive sensing emerges as a new sampling theory. Sampling and compression get on simultaneously in CS. In other words, it no longer deals with analog to digital conversion, but analog to information conversion (AIC), which dramatically reduces the sampling rate as well as the cost of data storage and transmission [1]. Efficient reconstruction method is the key in CS. There are three main types [2] as show in the following.

Convex relaxation. These methods only need a small number of observations, but with high computational burdensome and long time to carry out.

Natarajan Meghanathan et al. (Eds) : AIAA, WiMNeT, DPPR - 2018
pp. 09–21, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.81602

Combinatorial algorithms. These methods can support for rapid reconstruction. Nevertheless, the quality of reconstructed image is poor than that of convex relaxation.

Greedy pursuits. These algorithms are intermediate in their running time and reconstruction quality [3], which have simple structure and low computation. Examples include Matching Pursuit(MP)[4],Orthogonal Matching Pursuit(OMP)[5], regularized Orthogonal Matching Pursuit(ROMP)[6], Stagewise Orthogonal Matching Pursuit(StOMP)[7]. Wu J & Liu F[8] proposed the Multivariate Pursuit Algorithm(MPA), which is based on the multi-scale model to reconstruct the image in the wavelet domain. But it's dependent on the scale model too much. Budhiraja S [9] improves the OMP algorithm for better performance, but running time is still longer than the StOMP algorithm.

We should realize that the reconstruction time and the quality of reconstructed image are mutually restricted in CS. In addition, when CS is applied to 2D images, the measurement matrix can be fairly large, which makes several challenges including a computationally expensive reconstruction process and huge memory required to store the sampling operator. To solve this problem, Gan L proposed Block Compressed Sensing (BCS)[10]. It only needs to store the random matrix of blocked image that has compact size. And the encoder can transmit each block after its linear projection instead of waiting until the entire image is measured. However, the quality of the reconstructed image declines because BCS does not consider the global sparsity, and neglect the difference of the information distribution between blocks. Therefore, it is still a key problem to find the effective and reasonable method to segment and reconstruct images.

Therefore, this work focuses on how to design efficient image reconstruction method. The aim to attain better balance in the reconstruction time and the reconstruction quality, then we proposed the improved block stage wise regularized orthogonal matching pursuit (BStROMP) method. Simulation shows that the proposed algorithm can support for rapid reconstruction and ensure the quality of the images.

The rest of this paper is organized as follows. Section II provides the mathematical model of image reconstruction based on compressed sensing. Section III presents the proposed BStROMP algorithm. Section IV reports analysis of the simulation results followed by conclusions in Section V.

2. THE MATHEMATICAL MODEL OF IMAGE RECONSTRUCTION BASED ON CS

The framework of CS theory includes three aspects, the sparsity of the signal, the design of the measurement matrix and the robustness reconstruction method [1]. The idea of sparse expression is that when using the appropriate orthogonal basis as signal representation, the vast number of natural signals and images can be sparse or compressible[11]. The role of the measurement matrix is to reduce the dimensions thus CS can reduce the number of samples, and the reconstruction matrix needs to be irrelevant to the sparse representation matrix. The reconstruction method means that we can recover the original signal by solving the optimization problem [12]. In Figure 1, sampling and compressing are at the same time. Sampling M discrete data to avoid redundant samples directly, it reduces the cost of storage and transmission. What's more, the complexity of information acquisition shifts from the sampler to the receiver [13].

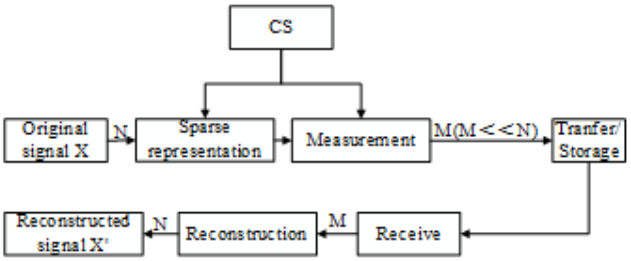


Figure 1. Flow chart of signal reconstruction based on compressive sensing

$x=[x(1),x(2),\ldots x(N)]\in R^N$. Sampling x at the rate much lower than Nyquist rate, the measurement matrix is denoted by $\Phi(M\times N,\ M<<N)$. Then we can get the $M\times 1$ measurement vector $y\in R^M$ from equation (1) and (2).

$$y=\Phi x \tag{1}$$

$$x=\psi\alpha \tag{2}$$

In equation (2), $\psi\in R^{N\times N}$, sparse vector $\alpha\in R^N$ is called the K-sparse expression of x[14] , which means that it contains only K nonzero element [11] . From equation (1) (2), we have

$$y=\Phi\psi\alpha=\Theta\alpha \tag{3}$$

Among equation (3), $\Theta=\Phi\psi$, it is called holographic dictionary or library[11]. When Θ satisfies the condition of RIP, we can obtain the estimation $\hat{\alpha}$ of the sparse signal α by equation (3). Then we can reconstruct the original signal $\hat{x}=\psi\hat{\alpha}$. CS theory uses optimization strategies that focus on finding the sparsest signal that matches with the M projections in y to reconstruct the signal. It is specific outstanding advantages. In other words, the reconstruction problem can be recast as a minimization l_0 problem.

$$\min_\alpha \|\alpha\|_0 \ s.t. y=\Phi\psi\alpha \tag{4}$$

Unfortunately, the l_0 optimization problem is NP-hard. Much work has been done to show that under some specific circumstances, a convex optimization problem can be used to recover such signals. These results show that the sparse recovery problem is equivalent to the l_1 convex program [11].

$$\min_\alpha \|\alpha\|_1 \ s.t. y=\Phi\psi\alpha \tag{5}$$

In fact, real signal would always accompany with noise. The equation (5) can convert to the optimization problem in equation (6) [11].

$$\min_\alpha \|\alpha\|_1 \ s.t. \|y-\Phi\psi\alpha\|_1\le\varepsilon \tag{6}$$

For a positive constant C, if $M \geq C\mu^2(\Phi,\psi)Klog(N/\delta)$ holds, the solution α of the l_1 norm minimization problem can be accurately calculated by the probability of 1- δ. Thereby the high-dimensional discrete vector x can be reconstructed from the low-dimensional sampling vector y with a probability of 1-δ[11].

3. THE BSTROMP ALGORITHM

Based on block CS, different reconstruction methods require different running time, and they would lead to different reconstruction quality. Basic pursuit (BP) [15] is a l_1 convex program, with high reconstruction accuracy and good reconstructed image quality. However, its running time is the longest among all the method that we mentioned. Compared with BP, the running time of the OMP [5] is reduced, and the quality of reconstructed image is relatively good, but the running time is still long. StOMP [7] reduces its running time dramatically, but the reconstructed image quality is worst. Budhiraja proposed modified orthogonal matching pursuit (MOMP) [16], the quality of reconstructed image improves a lot, but MOMP also need long time to run. Gao X presents Block Sparse Adaptive Regularized Matching Pursuit (BSARMP)[17]. To a certain extent, the running time is short than other methods, but the time is still too long. What's more, it does not have the same performance as BP.

In order to solve the problem above, we improve the StOMP algorithm, and then propose the BStROMP.

3.1. Block Compressed Sensing of Natural Images

Suppose x is a $N_r \times N_c$ image, We can transfer the 2-D image to 1-D signal, and then sample the 1-D signal to take M CS measurements with lower sample rate, and the original image can be recovered by reconstruction algorithm. However, when the length of signal is too large, the measurement matrix and the storage needed would be too large [10]. Therefore, we segment the image into several blocks as shown in figure 2.

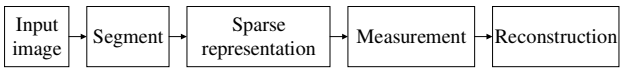


Figure 2. Process of block compressed sensing

The input image with $N=N_rN_c$ pixels in total will be divided into n small blocks with size of $B \times B$. Let $x_i, i = 1,2,...x_n (n = N/B^2)$ represent the vectorized signal of the i-th block through raster scanning. We use the same matrix Φ_B to sample each block[10]. The corresponding output of CS vector can be expressed as

$$y_i = \Phi_B x_i, i = 1,2,...,n \tag{7}$$

Where Φ_B is represented as Gaussian matrix with size of $n_B \times B^2$ generally, in which, $n_B = \left\lfloor \frac{MB^2}{N} \right\rfloor$, and B can be 8, 16 ,32 and so on. For the whole image, the equivalent sampling operator Φ is thus a block diagonal matrix taking the following form [10]

$$\Phi = \begin{vmatrix} \Phi_B & 0 & 0 & 0 \\ 0 & \Phi_B & 0 & 0 \\ 0 & 0 & \Phi_B & 0 \\ 0 & 0 & 0 & \Phi_B \end{vmatrix} \tag{8}$$

In equation (8), we can figure out that we never need to storage a full $n \times N$ matrix. Instead, Block CS is memory efficient for we just need to store a $n_B \times B^2$ Gaussian ensemble. In this way, we can reduce the computational complexity. However, if we use the same operator to sample each block, the loss of difference between blocks would cause the quality decline of the reconstructed image. What's more, it may cause block artifacts, ringing effect or distortion in edge detail. Therefore, according to the characteristics of the image, we divide all the image blocks into three types: Smooth, Edge and Texture. According to their different features, we set different weight for the number of measurement times. The image blocks are classified by calculating the variance

$$v = \frac{1}{B \times B} \sum_{j=1}^{B \times B} (\mathrm{p_j} - \overline{\mathrm{p}})^2, \overline{\mathrm{p}} = \frac{1}{B \times B} \sum_{j=1}^{B \times B} \mathrm{p_j} \tag{9}$$

$$v' = \frac{v - v_{\min}}{v_{\max} - v_{\min}} \tag{10}$$

Where v is the variance, and p_j is the j-th pixel value of the block. $v_{\max}$ and $v_{\min}$ are the maximum and the minimum of the variance in all the image blocks respectively. By equation (10), we can classify X.

$$x_i = \begin{cases} Smooth, if \quad v' \le T \\ Texture, if \quad v' > T_2 \\ Edge, \ \ if\ T_1 < v' \le T_2 \end{cases} \tag{11}$$

When v'$\leqslant$0.1 , x_i is regarded as smooth , and when v'$\geqslant$0.3 is classified as Texture , and the rest is Edge.

Assume that we have n image blocks in total, we set the sampling rate of smooth, edge, texture to r_1, r_2, r_3 respectively. And the number of blocks of each part is $n_1, \mathrm{n}_2, \mathrm{n}_3$ respectively.

Suppose r is the fixed sampling rate for the whole image, we have

$$n \times B^2 \times r = n_1 \times B^2 \times r_1 + n_2 \times B^2 \times r_2 + n_3 \times B^2 \times r_3 \tag{12}$$

In equation (12), we set $r_1 = R, r_2 = 3R, r_3 = 5R$. Then we have

$$R = \frac{N \times r}{N_1 + 3N_2 + 5N_3} \tag{13}$$

We can calculate r_1, r_2 and r_3 by R. We control the sampling rate of different part through different weight. Higher sampling rates for images containing rich information prevent the loss of details and texture information after refactoring. The lower sampling rate for less information can also lead to good performance.

3.2. ACGP algorithm

In StOMP, we calculate the correlation coefficient first and then use a hard threshold to judge the new column to update the support set. In this way, we can speed up the iteration but the support set updated is not optimal in each iteration, which makes the reconstruction accuracy greatly reduced [3]. In order to solve the problem, we use regularization method to verify the threshold. To reduce the complexity of the StOMP algorithm, we adopt an approximate conjugate gradient algorithm to estimate the approximate solution the signal.

Set the equation

$$Ax = b, x \in R^n \tag{14}$$

Where $A \in R^{m \times n}$ is a symmetrical positive determined matrix. We use conjugate gradient algorithm[18] to solve equation (14), it is equivalent to find the minimum of y in equation (15), i.e. to solve a quadratic optimization problem.

$$y = \frac{1}{2} x^T Ax - b^T x, x \in R^n \tag{15}$$

Assume that $p, q \in R^n$, both p and q are not zero. If equation (16) holds, then the vectors p and q are A orthogonal or conjugate.

$$p^T Aq = 0 \tag{16}$$

if $d_1, d_2, \ldots, d_n \in R^n$, for $\forall i \neq j$ and equation (17) holds, then $D = \left\{ d^i, d^2, ..., d^n \right\}$ is the conjugate vector group of A.

$$(d^i)^T Ad^j = 0 \tag{17}$$

If the orthogonal matrix A is a unit matrix, then p and q are orthogonal obviously, and we can see that the conjugate is the expansion of orthogonal [11].

For $\forall x^0 \in R^n$, after n iterations at most, we can get the solution of equation (14) by equation (18).

$$\hat{x}^n = \hat{x}^{n-1} + \alpha^n d^n \tag{18}$$

Where d is an update direction or conjugate. We adopt the approximate Conjugate Gradient Pursuit (ACGP) method for lower computational complexity and less storage space. We don't need all previously chosen directions but the current to calculate the new direction. The progress is as follows.

Without loss of generality, we set b0=1, and the gram matrix $G=\Phi^T\Phi$, and g^n is the product of the measurement matrix and residual in n-th iteration. Let G conjugate to all the previously chosen directions, then we have

$$< Gd^{(n-1)}, g^n + b_1 d^{(n-1)} >= 0 \tag{19}$$

We can calculate b_1 from equation (19).

$$b_1 = \frac{< \phi d^{n-1}, \Phi g^n >}{\left\| \Phi d^{n-1} \right\|^2} \tag{20}$$

The new update direction is

$$d^n = g + d^{n-1} b_1 \tag{21}$$

With step sizes

$$\alpha^n = \frac{< r^n, d^n >}{< d^n, A d^n >} \tag{22}$$

In summary, we can obtain the residual after this iteration

$$r^n = b - Ax^n = r^{n-1} - \alpha^n A d^n \tag{23}$$

3.3. The framework of BStROMP

Based on the mathematical model in part 2 and the improved algorithm in 3.1 and 3.2, the framework of this paper is shown in figure.3.

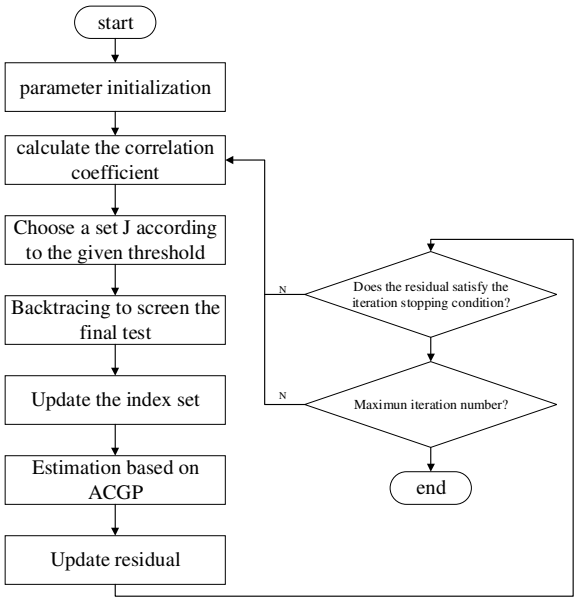


Figure 3. Process of BStROMP

Given the $M \times N$ measurement matrix A=Φψ, and the observation vector y with size of M×1. Through the proposed method in this paper, we can reconstruct the original signal to get the best estimation $\hat{x}$. We will discuss the detail as follows.

(1) Initialization. Let r_0=y, $\Lambda_0=\emptyset$, t=1, where Λ_t is the index in t-th iteration. A_0 is the column vector collection of the matrix A selected by the index.

(2) Calculate. Calculate the correlation coefficient (the absolute value of product) $C_t = \left| \left[A^T r_{t-1} \right] \right|$.

(3) Identify. Initial selection $J_t = \{ j : | c_t(j) > T_h = t_s \left\| r_t \right\|_2 / \sqrt{M} | \}$, where T_h is the hard threshold coefficient.

(4) Secondary choosing. If $\| J_t \|_0 \neq 0$, then we use regularized principle $|c(i)| \leqslant 2|c(j)|, i,j \in J_t$ to choose a set $\left\| u |_{J_0} \right\|_2$ with maximum energy ; on the contrary, let J_t=max u_i, i=1,…,N.

(5) Update the support set. $\Lambda_t = \Lambda_{t-1} \cup J_t, A_t = A_{t-1} \cup \lambda_j \left(j \in J_t \right)$, λ_j is the j-th column of A

(6) Calculate $g^t = A_{\Lambda^t}^T r^{t-1}$. The step size $\alpha_t = < r^{t-1}, A_{\Lambda^t} d_{\Lambda^t} >$, update the new direction d_{Λ^t}.

(7) The estimation of original signal $\hat{x}_{\Lambda^t} = \hat{x}_{\Lambda^t}^{t-1} + \alpha^t d_{\Lambda^t}$.

(8) Update the residual $r_t = y - A_t \hat{x}_t$.

(9) t=t+1. If t$\leqslant$S, then go back to (2). If the residual is equal to zero or $\hat{x} = x_t$, r $= r_t$, then quit the algorithm.

(10) Output $\hat{x}$.

The coefficient t_s in (3) usually rages from 2 to 3. The setting of the threshold is very suitable for the reconstruction of Gaussian random signal, but it cannot guarantee the reliability and effectiveness for non-Gaussian signal. When algorithm executes, the atom chosen by the threshold may not meet the conditions, resulting that the support set cannot be optimal in the subsequent cycle. Therefore, we use the regularization to verify the reliability of the threshold.

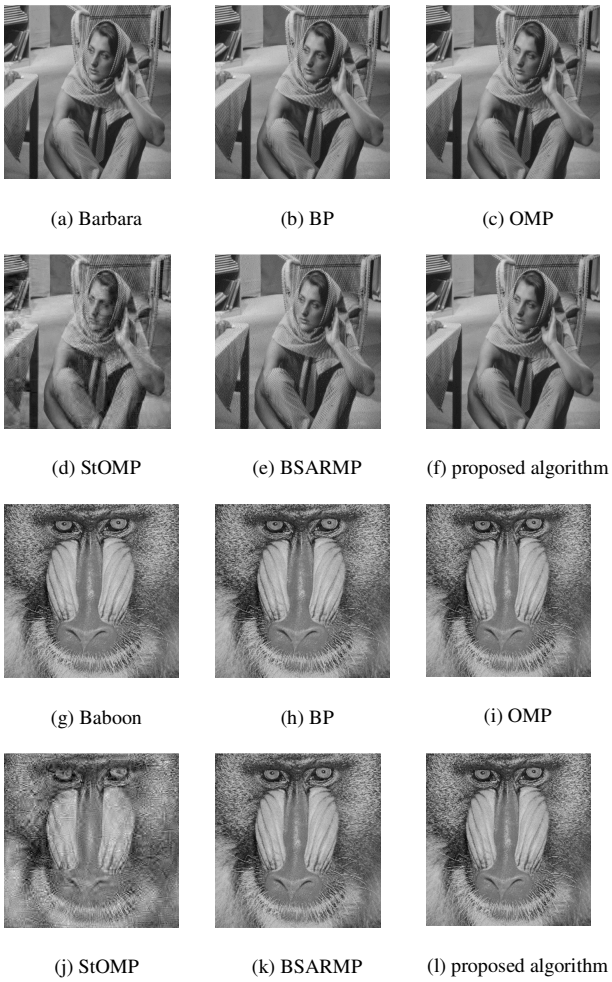
If the number of atoms selected in step (3) is greater than 0, the threshold is reasonable, and the selected atoms are regularized to select the optimal set of atoms. If step (3) cannot chose the atom we need, the atom corresponding to the maximum correlation coefficient in the step (2) is chosen. In this case, there is only one atom chosen, and no regularization is necessary.

Step (6) and (7) adopt the ACGP algorithm to solve the estimation of the signal instead of the least squares method. Because of the definite update direction, the speed of iterative convergence is accelerated. Based on the regularization idea above, the performance of BStROMP is improved a lot.

4. SIMULATIONS

The main criteria to evaluate the performance of CS reconstruction methods are the running time and the reconstructed image quality. Although some reconstruction algorithm have short running time, the quality of the reconstructed image is poor and some of them need long time to run, but the quality of reconstructed image is good. The two criteria restrict each other; we cannot guarantee that the time and the quality are both optimal. We choose Barbara and Baboon in our

experiments to reconstruct them by BP, OMP, StOMP, BSARMP and proposed algorithm respectively. Then we can compare the running time and the reconstructed image quality of those algorithms. The peak signal-to-noise ratio (PSNR) and reconstruction error are used as objective criteria to evaluate the quality of reconstructed image. In the experiment, the overall sampling rate is R=0.5, R=M/N, where M is the total number of samples, and N is the total pixels of the original image. The simulation results are shown below.



(a) Barbara (b) BP (c) OMP

(d) StOMP (e) BSARMP (f) proposed algorithm

(g) Baboon (h) BP (i) OMP

(j) StOMP (k) BSARMP (l) proposed algorithm

Figure 4. Comparison with other CS techniques

Subjectively, according to the analysis of the simulation results of the image above, we can see that the image obtained by the BP algorithm is the best, and the image quality of StOMP is the worst. The performance of proposed algorithm is slightly worse than the BP algorithm, but the difference between all the images above is relatively insignificant from human visual.

To make further compare the performance of each algorithm, we obtain the mean of the evaluation parameters after 50 simulation experiments, as shown in the follows.

Table 1. Performance comparison of different algorithms

Algorithm		BP	StOMP	OMP	BSARMP	proposed algorithm
PSNR(dB)	Barbara	35.460	28.104	31.142	32.539	33.535
	Baboon	32.383	27.818	29.738	30.106	31.379
Time (s)	Barbara	15.684.	0.841	8.698	2.564	1.783
	Baboon	15.924	0.892	9.015	2.596	1.836
Relative Error	Barbara	0.0235	0.3614	0.0483	0.0394	0.0334
	Baboon	0.0151	0.380	0.0492	0.0410	0.0322

In order to better compare the performance of each algorithm, we also draw the graph of Barbara as follows.

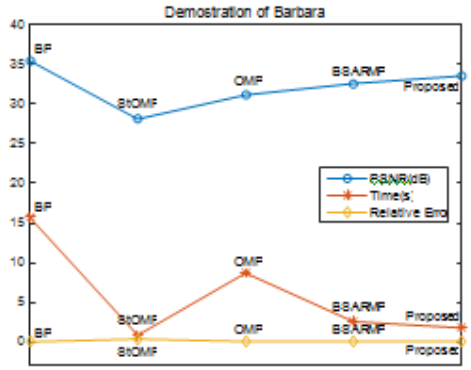


Figure 5. Demostration of Barbara with different CS methods

As seen from table 1 and figure 5, the PSNR of BP is the highest and its quality of reconstructed image is the best when the same sampling rate is used. However, it's running time is the longest among all algorithms. StOMP algorithm has the lowest PSNR, the worst reconstructed image quality, but its running time is the shortest. Although the PSNR and the relative error of the proposed algorithm are lower than BP, compared to other algorithms, the proposed algorithm have better performance, and the running time is much lower than BP, which is only slower than that of StOMP. It can be seen that the time complexity of the proposed algorithm is only higher than that of StOMP, but it is better than BP, OMP and BSARMP. Moreover, the proposed algorithm is improved based on StOMP, with the idea of regularization. Although the reconstruction time is longer than StOMP, the quality of reconstructed image has been greatly improved.

The size of block is also an important factor. We set different size to test the performance of the proposed algorithm: 8×8, 16×16, 32×32. Chose Barbara, compare the PSNR in each simulation at different sampling rate. The results are shown in Table 2.

Table 2. PSNR of different block size

R	8×8	16×16	32×32
0.1	20.034(dB)	22.612(dB)	24.069(dB)
0.2	22.581(dB)	24.948(dB)	27.044(dB)
0.3	26.249(dB)	27.22(dB)	29.76(dB)
0.4	29.452(dB)	31.156(dB)	32.630(dB)
0.5	31.156(dB)	32.889(dB)	33.535(dB)

For the typical image such as Lena (256×256), use the image block, and calculate the peak signal-to-noise ratio (PSNR) at different sampling rates, the unit is dB, and the experimental results are shown in Figure 6.

As can be seen from table 2, the larger the block image, the better the reconstructed accuracy. However, the storage required for sampling and the computational complexity for reconstruction are also increased. For the representative image such as Lena (256×256), with the block size of 32×32, the PSNR at different sampling rate is shown in Figure 6.

Figure 6. Demonstration of BStROMP with the same block size(32×32) but different sampling rate

As can be seen from the figure above, with the sampling rate increases, that is, the number of measurements increases, the amount of information obtained as well as the precision would increases. The original image can be reconstructed with high precision when certain conditions are satisfied. Nevertheless, we need to deal with more data at the same time; the complexity and the running time would also increase.

In summary, compared with StOMP, the proposed algorithm improves the reconstructed quality, although the running time is slightly longer than StOMP, which is shorter than other algorithms. The reconstructed quality of the proposed algorithm is inferior compared to BP, but better than OMP and BSARMP.

5. CONCLUSIONS

The BStROMP algorithm proposed in this paper reduces the storage and computational complexity required in encoder by using the idea of block CS. Using the atom selection method in StOMP, which is restricted by regularization, improves the precision of reconstructed image. We implement fast reconstruction as well as improving the reconstructed image quality in decoder. The simulation results show the robustness of the proposed algorithm.

However, the proposed algorithm cannot reconstruct the original image well when the sampling rate is extremely low. Therefore, we can further to find a better observation matrix or use other sparse transform to optimize the sparse representation of images.

The CS reconstruction methods have two problems to resolve, the reconstruction time and the quality of reconstructed image. However, they are contradictory; we cannot guarantee both in the meantime. In addition, the quality of reconstructed image would also affected by environment, so we plan to deal the reconstructed image with post-processing enhancement.

ACKNOWLEDGEMENTS

This work is supported by National Nature Science Foundation under Grant 61473322 and 61772140, the Technology Planning Project of Guangdong Province under Grant 2017A010101021, the Guangdong Natural Science Foundation - doctoral program under Grant 2016A030310335, and the Foshan science and technology innovation project under Grand 2016AG101823 and 2016AG101793.

REFERENCES

[1] Shi Guang-Ming, Liu Dan-Hua & Gao Da-Hua, (2009) “Compressed Sensing Theory and Its Research Progress”, Journal of Electronics, Vol. 37, No. 5, pp1070-1081.

[2] Li Shen, Ma Cai-Wen & Li Yan, (2013)"Review of Compressed Sensing Reconstruction Methods", Infrared and Laser Engineering, Vol. 42, No. S01, pp225-232.

[3] Yang Zhen-zhen, Yang zhen & Sun Lin-Hui, (2013)"A Review of Orthogonal Matching Pursuit Algorithms for Signal Compression Reconstruction ", Signal Processing, Vol. 29, No. 4, pp486-496.

[4] Mallat S G & Zhang Z, (1993)"Matching pursuits with time-frequency dictionaries", IEEE Transactions on Signal Processing, Vol. 41, No. 12, pp3397-3415.

[5] Tropp J A & Gilbert A C, (2007)"Signal recovery from random measurements via orthogonal matching pursuit", IIEEE Transactions on Information Theory, Vol. 53, No. 12, pp4655-4666.

[6] Needell D & Vershynin R, (2009)"Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit", Foundations of computational mathematics, Vol. 9, No. 3, pp317-334.

[7] Donoho D L, Tsaig Y & Drori I, (2012)"Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit", IEEE Transactions on Information Theory, Vol. 58, No. 2, pp1094-1121.

[8] Wu J, Liu F& Jiao L C, (2011)"Multivariate compressive sensing for image reconstruction in the wavelet domain: using scale mixture models", IEEE Transactions on Image Processing, Vol. 20, No. 12, pp3483-3494.

[9] Budhiraja S, (2015)"Image reconstruction using modified orthogonal matching pursuit and compressive sensing", 2015 IEEE International Conference on Computing, Communication & Automation (ICCCA), Vol. , No. , pp1073-1078.

[10] Gan L, (2007)"Block compressed sensing of natural images", 2007 15th IEEE International Conference on Digital Signal Processing, Vol. , No. , pp403-406.

[11] Zhang Xian-Da (2013) Matrix Analysis and Application, Beijing: Tsinghua University Press.

[12] Wenze S & Zhihui W, (2012)"Advances and perspectives on compressed sensing theory", Journal of Image and Graphics, Vol. 17, No. 1, pp1-12.

[13] Li Zhi-Lin, Chen Hou-Jin & Yao Chang, (2012)"An Improved Multiscale Compressed Sensing Scheme ", Photoelectron Laser , Vol. 23, No. 7, pp1403-1410.

[14] Candes E & Romberg J, (2007)" Sparsity and incoherence in compressive sampling", Inverse problems, Vol. 23, No. 3, pp969.

[15] Chen S S, Donoho D L & Saunders M A, (2001)"Atomic decomposition by basis pursuit", SIAM review, Vol. 43, No. 1, pp129-159.

[16] Budhiraja S, (2015)"Image reconstruction using modified orthogonal matching pursuit and compressive sensing", 2015 IEEE International Conference on Computing, Communication & Automation (ICCCA), Vol. , No. , pp1073-1078.

[17] Wang A, Gao X & Gao Y, (2014)"A Modified Image Reconstruction Algorithm Based on Compressed Sensing", 2014 IEEE Fourth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC), Vol. , No. , pp624-627.

[18] Hestenes M R & Stiefel E (1952) Methods of conjugate gradients for solving linear systems, NBS.

AUTHORS

Xiong-Yong Zhu, Ph.D , Lecturer in the Department of Computer Science, Guangdong University of Education. His research interest covers digital image processing, video signal processing, computer vision.

Shun-Dao Xie, Ph.D student at the School of Electronics and Information Technology, Sun Yat-Sen University. His research interest covers digital image processing, digital integrated circuit design, and Internet of things.

Guo-Min Chen, Received the M.S. and Ph.D. degrees from School of Information Science and Technology, Sun Yat-sen University, China, in 2003 and 2009. He is an associate professor of Guangdong University of Education. His areas of interest include data mining, machine learning, pattern recognition, and image processing.

Wen-Fang Wu, Master student at the School of Electronics and Information Technology, Sun Yat-Sen University. Her main research interest is digital image processing.

Hong-Zhou Tan, Professor at the School of Electronics and Information Technology, Sun Yat-Sen University. His research interest covers broadband communications, signal processing, identification and modeling of complex systems, and semicondector IC design. Corresponding author of this paper.

INTENTIONAL BLANK

TIME-DOMAIN SIGNAL MANAGEMENT FOR OFDM SIGNALS

Takuya Kazama[1], Kazuki Miyazawa[2], and Masahiro Muraguchi[3]

[1,2]Faculty of Engineering, Tokyo University of Science, Tokyo, Japan
[3]Tokyo University of Science, Tokyo, Japan

ABSTRACT

We have found out that the CAZAC- OFDM accords the amplitude of IFFT output signal with the amplitude of input DATA and the time ordering of IFFT output signal is unambiguously determined. That is, the OFDM time-domain signals, which are composed of many sinewaves, can be shaped by CAZAC precoder. As one application example that can use this characteristic of CAZAC precoder, we propose a new technique of symbol timing estimation, which enable to avoid the use of the preambles and guard-intervals. Conventional OFDM systems introduce guard-intervals for symbol timing estimation and reduction of multipath effect. Visible light communications (VLCs), which are one kind of line-of–sight communications, does not require consideration of the multipath channel. Therefore, if we embed null data at a fixed position in time-domain, we will easily estimate the symbol timing in the receiver side.

KEYWORDS

OFDM; CAZAC sequence; Zadoff-Chu sequence; Symbol Timing Estimation; VLC.

1. INTRODUCTION

In recent years, with the development of wireless communication technology, the usage of wireless communication by smart phones, tablets, laptops, etc. has been increasing steadily. Meanwhile radio frequency bands are becoming exhausted, and Visible Light Communication (VLC) having much shorter wavelength than radio waves is being studied. Among visible light communication, studies using light emitting diodes (LED) become increasingly active [1], [2]. The reason for this is that high-speed communication due to the development of high-speed responsiveness of LEDs has become possible [3], and ease of dissemination due to the proliferation of illumination using LEDs also in general household.

A major drawback of VLC is band-limited operation due to the frequency characteristics of LEDs. Therefore, VLC requires a modulation scheme with high spectral efficiency. Orthogonal frequency division multiplexing (OFDM) scheme is very suitable for the VLC, because its great advantage of spectrum utilization that is about twice spectral efficiency compared with single carrier modulation scheme.

The OFDM system, however, requires highly exact timing for the symbol synchronization. The symbol timing shift in the time-domain on the receiver incurs a serious distortion after demodulation.

Therefore, conventional OFDM systems introduce preambles and guard-intervals for symbol timing estimation. Although the preamble enables high-precision symbol synchronization, it is

Natarajan Meghanathan et al. (Eds) : AIAA, WiMNeT, DPPR - 2018
pp. 23–35, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.81603

only applied intermittently in a communication in which the beginning of a signal does not exist, for example, a signal continuously transmitted such as a broadcast. Even though the guard interval enables synchronization for each symbol, a significant throughput reduction is drawn, and it is wasteful considering the application in the VLC which is a line-of-sight wireless system anddoes not need to consider the multipath signals.

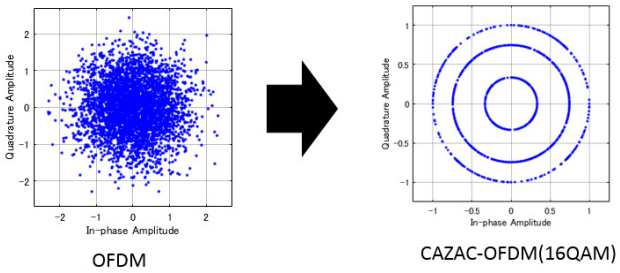


Figure1.Time domain Signal without and with CAZAC precoder

Another well-known drawback of OFDM signal is its high peak-to-average power ratio (PAPR). As OFDM signal is essentially a sum of multiple subcarrier signals aligned in frequency-domain, its probability density function in time-domain resembles Gaussian distribution, and thus its amplitude has high PAPR. The constant amplitude zero auto-correlation (CAZAC) precoding had been proposed as one of PAPR reduction techniques for OFDM signals [5]. Afterwards we had found out that the CAZAC precoding (CAZAC-OFDM) made the PAPR of M-array quadrature amplitude modulation (M-QAM) OFDM signals into the PAPR of M-QAM single-carrier signals [6].Moreover, we had demonstrated the theoretical verification of the CAZAC-OFDM [7]. Here, we also had found out the CAZAC- OFDM accorded the amplitude of IFFT output signal with the amplitude of input DATA and the time ordering of IFFT output signal was unambiguously determined. We verify that this determination is caused by compatibility with the IFFT operation, not by the autocorrelation properties of the CAZAC sequence. Moreover, if we embed null data at a fixed position in time-domain, we will easily detect the symbol timing shift in the receiver side.

The rest of the paper is outlined as follows: Section II contains overview of symbol timing estimation in conventional OFDM system and verifies that time domain signals of OFDM which cannot be manipulated in conventional can be shaped by using CAZAC precoder, and Section III contains description of proposed system using the characteristics of CAZAC-OFDM, and section IV presents performance evaluation and computer simulation results, and section V concludes the paper.

2. VISIBLE LIGHT COMMUNICATION AND CAZAC SEQUENCE

2.1 Visible Light Communication (VLC)

Visible light communication has developed rapidly as an attractive technology for supplementing conventional radio frequency communication, which is made possible by the rise of LEDs. One of the important characteristics of the reason why LEDs is used for visible light communication is fast responsiveness.

LEDs are able to follow frequency high enough to send useful data and it is a frequency that is too high to be perceived by the human eye. Therefore, it is possible to complement the role of the LEDs widely used as the lighting and the function as the communication device.

2.2 OFDM system

As well-known, OFDM system enables the limited frequency bandwidth of LEDs or LDs can be fully-used.

In an OFDM transmitter, an inverse fast Fourier transform (IFFT) is performed after primary modulation of a data sequence. As a result, the value after primary modulation is placed on each subcarrier, and a spectrum is formed in which each subcarrier overlaps. Even if each subcarrier overlaps, due to mutual orthogonality, respectively demodulation is possible, so that frequency utilization efficiency is excellent.

The nth sample with the input signal $\boldsymbol{X}$ (length N) after mapping is defined as

$$x[n] = \frac{1}{N}\sum_{k=0}^{N-1} X[k]e^{j\frac{2\pi}{N}kn} \qquad (1)$$

Where $X[k]$ is the frequency-domain signal, and N is the number of subcarrier, $j = \sqrt{-1}$. After IFFT, a guard interval is inserted in the baseband signal. The OFDM symbol is generated by the above procedure.

Since OFDM signals are composed of signals on a symbol-by-symbol basis, when performing FFT for demodulation, it is necessary to find the beginning of the symbol and open the FFT window at an exact timing.

Generally, symbol timing is estimated by using correlation characteristics with known or existing signals [8], [9]. One method is to detect a peak by cross correlation using a preamble. It adds a known sequence to the beginning of the signal and establishes symbol synchronization by calculating the correlation value between the known sequence and the signal.

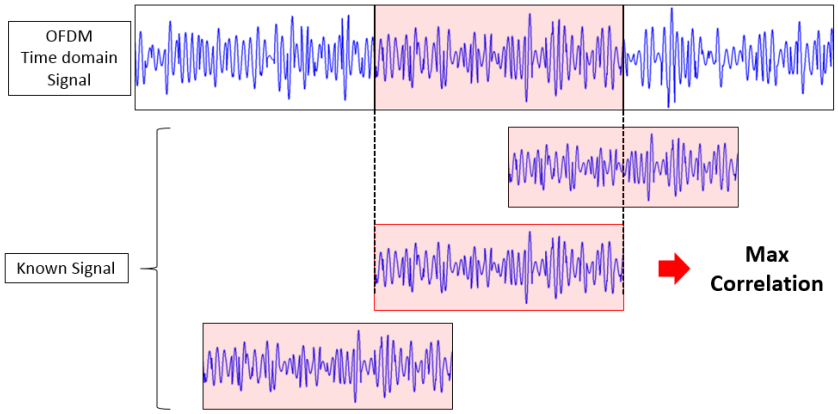


Figure 2.Symbol timing estimation by cross-correlation

When the known sequence matches the signal, the correlation value becomes the maximum, so we use this to detect the beginning of the symbol.

Another, there is estimation using the guard interval. In this case, by adding a guard interval, the same value exists in the portion apart from the copied portion by the original symbol length. In other words, the correlation between the normal waveform and the waveform delayed by the

original symbol length becomes higher in the copied part. From this, the symbol timing is estimated by detecting the peak of the correlation value.

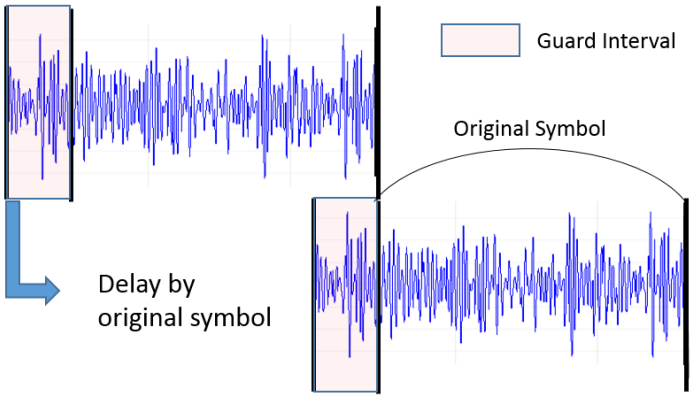


Figure 3. Symbol timing estimation using Guard Interval

2.3 Effect of Symbol Timing Offset (STO)

Consider the effect of STO when guard interval to be described later is used. First, FFT is defined as

$$\mathrm{X[n]} = \sum_{k=0}^{N-1} x[k] e^{-j\frac{2\pi}{N}kn} \qquad (2)$$

As can be seen from (1), (2), demodulation at the receiver cannot be demodulated without distortion unless all 0 ~ k – 1 samples are available. Here, when δ sample offset occurs in the time domain, assuming that inter-symbol interference does not occur, following as

$$\mathrm{X'[n]} = \sum_{k=0}^{N-1} x[k+\delta] e^{-j\frac{2\pi}{N}kn} \qquad (3)$$

Similarly, considering the correspondence of (1), (3), it becomes as follows

$$\begin{aligned} \mathrm{X'[n]} &= \frac{1}{\mathrm{N}} \sum_{k=0}^{N-1} \left\{ \sum_{m=0}^{N-1} X[m] e^{j\frac{2\pi}{N}m(n+\delta)} \right\} e^{-j\frac{2\pi}{N}kn} \\ &= \frac{1}{\mathrm{N}} \sum_{k=0}^{N-1} X[m] e^{j\frac{2\pi}{N}m\delta} \sum_{k=0}^{N-1} e^{j\frac{2\pi}{N}(m-k)n} \\ &= \mathrm{X[n]} e^{j\frac{2\pi}{N}n\delta} \qquad (4) \end{aligned}$$

From this, it can be seen that the STO within the guard interval generates a phase rotation proportional to the offset after the demodulation. The constellation when setting the FFT window such as Figure.4 is shown as Figure.5.

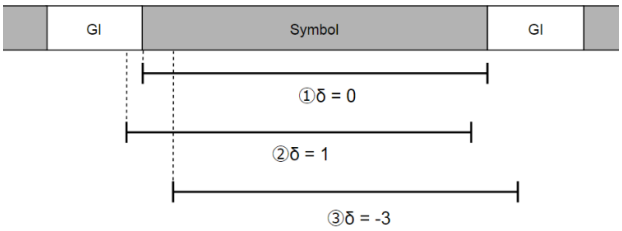


Figure 4.Symbol Timing Offset with OFDM

In the case of ①,the FFT window is set at an exact timing, so there is no offset, and a clean constellation appears after demodulation. In the case of ②, the FFT window is set before the exact point. Due to a guard interval, symbols are retrieved without interfering with adjacent symbols. However, since the extracted samples are cyclically shifted in the correct arrangement, the phase rotation proportional to the offset is added to the demodulated signal as shown in (4).It can be corrected by using the pilot signal in the symbol. In the case of ③, the FFT window is set after the exact point. In this case, it interferes with the adjacent symbol, not only the above-mentioned phase rotation but also distortion in the amplitude direction is added. Therefore, the signal deteriorates.

In conventional OFDM system, it is possible to correctly reproduce by correcting the offset within the guard interval by using the guard interval, but when the guard interval is not used, it is impossible to correct deterioration due to the offset. Therefore, symbol timing needs to be estimate exactly.

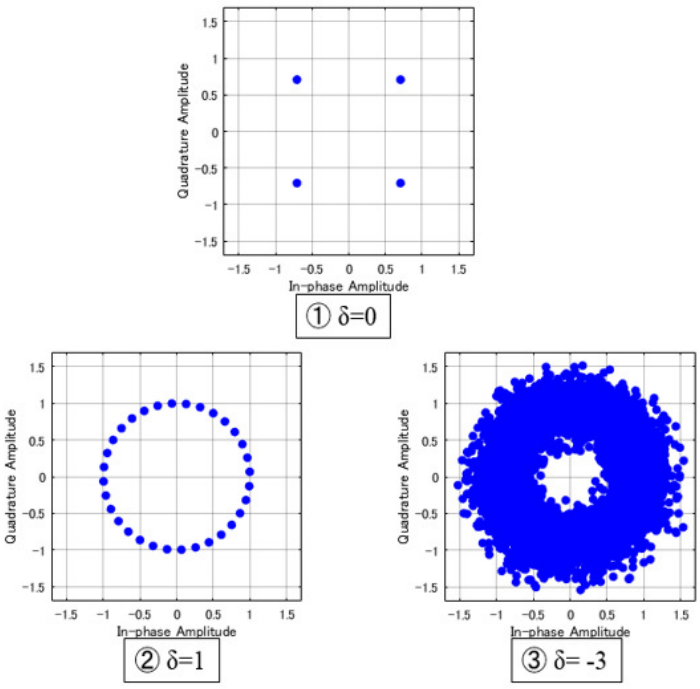


Figure 5. Constellation for ① ~ ③ Symbol Timing Offset with OFDM
(QPSK, FFT size:64)

2.4 CAZAC-OFDM

CAZAC-OFDM is an OFDM method precoded by using the CAZAC sequence shown in equation (5).

$$\boldsymbol{C}(k) = \begin{cases} \exp\left(j\frac{\pi u}{N}k^2\right) \text{ N is even} \\ \exp\left(j\frac{\pi u}{N}k(k+1)\right) \text{ N is odd} \end{cases} \quad k = 0,1,\dots,N-1 \tag{5}$$

In this case, when CAZAC precoder is applied in OFDM, the signal length N mostly takes a power of 2 because of FFT. Also, if u = 1, the Zadoff-Chu sequence used from the formula (5) is as follows.

$$\boldsymbol{C}(k) = \exp\left(j\frac{\pi}{N}k^2\right)(k = 0,1,\dots,N-1) \tag{6}$$

From (6), an N × N square matrix **M** is generated. The matrix equation is defined as

$$\mathbf{M} = \frac{1}{\sqrt{N}}\begin{bmatrix} \boldsymbol{c}_0 & \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_{N-1} \\ \boldsymbol{c}_N & \boldsymbol{c}_{N+1} & \cdots & \boldsymbol{c}_{2(N-1)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{c}_{(N-1)N} & \boldsymbol{c}_{(N-1)N+1\dots} & & \boldsymbol{c}_{(N-1)^2} \end{bmatrix} \tag{7}$$

In transmitter side, CAZAC precoder signal is generated by multiplying the mapping data by this matrix **M**. In receiver side, multiply the received symbol Y in frequency domain and the inverse matrix $\{\mathbf{M}^{\mathrm{H}}\}^{\mathrm{T}}$ after FFT.

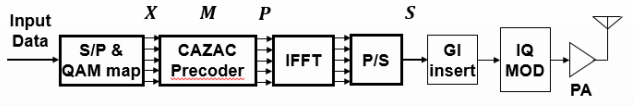


Figure 6. The receiver configuration using CAZAC precoder.

According to Figure.6, **X** is QAM Vector, **P** is Modulation Vector for IFFT, **M** is CAZAC Matrix.Then **S**as

$$\begin{aligned} \boldsymbol{S}(t_n) &= \frac{1}{N}\sum_{k=0}^{N-1}\boldsymbol{P}_k e^{j2\pi kn/N} \\ &= \sum_{k=0}^{N-1}\left\{\sum_{m=0}^{N-1} e^{j\pi(m+kN)^2/L}\boldsymbol{X}_m\right\} e^{j2\pi kn/N} \\ &= \sum_{m=0}^{N-1} e^{\frac{j\pi m^2}{N^2}}\boldsymbol{X}_m\left\{\sum_{k=0}^{N-1} e^{\frac{j2\pi k(m+n)}{N}} e^{j\pi k^2}\right\} \end{aligned} \tag{8}$$

Where k is an integer not less than 0, so the following equation is developed.

$$exp(j\pi k^2) = \begin{cases} 1 & (k : \text{even}) \\ -1 & (k : \text{odd}) \end{cases} \tag{9}$$

To lead the equation (10) from the equation (9).

$$\exp(j\pi k^2) = (-1)^k \qquad (10)$$

Substituting the equation (10) into the equation (8) lead to the equation (11).

$$\boldsymbol{S}(t_n) = \sum_{m=0}^{N-1} e^{j\pi m^2/N^2} \boldsymbol{X}_m \left\{ \sum_{k=0}^{N-1} \left\{ -e^{j2\pi(m+n)/N} \right\}^k \right\} \qquad (11)$$

The inside of { } in equation (11) is the sum of the geometric progression. Therefore, equation (12) is derived.

$$\sum_{k=0}^{N-1} \left\{ -e^{j2\pi(m+n)/N} \right\}^k = \begin{cases} N & \left(-e^{j2\pi(m+n)/N} = 1 \right) \\ 0 & \left(-e^{j2\pi(m+n)/N} \neq 1 \right) \end{cases} \qquad (12)$$

When $2\pi(m+n)/N = 1$, $2(m+n)/N$ is an integer and odd number. Also, n and m are$0 \leq n \leq N-1$, $0 \leq m \leq N-1$. Equation (13) follows on account of these relationships.

$$m = \frac{\mathrm{N}}{2} - n \ (mod\ N) \qquad (13)$$

Equation (14) is derived from equations (11), (12) and (13).

$$\boldsymbol{S}(t_n) = \boldsymbol{X}\left(\left(\frac{N}{2} - n \right)_{mod\ N} \right) \boldsymbol{C}\left(\left(\frac{N}{2} - n \right)_{mod\ N} \right) \qquad (14)$$

From the above, we have theoretically derived $\mathbf{S}(\mathrm{t_n})$ for arbitrary n. The time waveform of CAZAC-OFDM is obtained by phase-rotating the value after mapping.

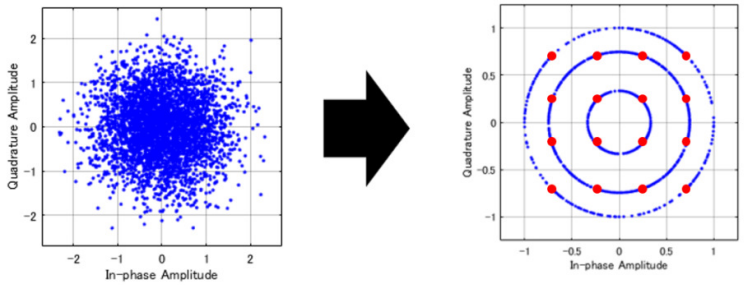


Figure7. Time domain Signals with CAZAC precoder influenced by the value after mapping.

Here, concrete calculation in the case of N = 4 is performed to understand well. According to Figure.6,$\mathbf{P} = \mathbf{MX}$is obvious. Therefore,$\mathbf{P}$ as

$$\begin{bmatrix} \boldsymbol{P}_0 \\ \boldsymbol{P}_1 \\ \boldsymbol{P}_2 \\ \boldsymbol{P}_3 \end{bmatrix} = \begin{bmatrix} 1 & e^{j\frac{\pi}{16}} & e^{j\frac{\pi}{4}} & e^{j\frac{9\pi}{16}} \\ -1 & -e^{j\frac{9\pi}{16}} & e^{j\frac{\pi}{4}} & -e^{j\frac{\pi}{16}} \\ 1 & -e^{j\frac{\pi}{16}} & e^{j\frac{\pi}{4}} & -e^{j\frac{9\pi}{16}} \\ -1 & e^{j\frac{9\pi}{16}} & e^{j\frac{\pi}{4}} & e^{j\frac{\pi}{16}} \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_0 \\ \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \\ \boldsymbol{X}_3 \end{bmatrix}$$

$$= \begin{bmatrix} X_0 + e^{j\frac{\pi}{16}}X_1 + e^{j\frac{\pi}{4}}X_2 + e^{j\frac{9\pi}{16}}X_3 \\ -X_0 - e^{j\frac{9\pi}{16}}X_1 + e^{j\frac{\pi}{4}}X_2 - e^{j\frac{\pi}{16}}X_3 \\ X_0 - e^{j\frac{\pi}{16}}X_1 + e^{j\frac{\pi}{4}}X_2 - e^{j\frac{9\pi}{16}}X_3 \\ -X_0 + e^{j\frac{9\pi}{16}}X_1 + e^{j\frac{\pi}{4}}X_2 + e^{j\frac{\pi}{16}}X_3 \end{bmatrix} \quad (15)$$

S is the IFFT-operated signal of **M**. Therefore $\mathbf{P}_0 \sim \mathbf{P}_3$ is respectively multiplied by following phase-rotation value for subcarriers of $f_0, 2f_0, 3f_0, 4f_0$ and summed.

$$e^{j\frac{2\pi(n+1)}{4}t_m} \quad (n=0,1,2,3) \quad (16)$$

t_m: IFFT output time order

$(t_m = 0,1,2,3)$

First, calculate $\mathbf{S}(t_0)$.

$$\begin{aligned} \mathbf{P}_0 e^{j0} &= \mathbf{X}_0 + e^{j\frac{9\pi}{16}}\mathbf{X}_1 + e^{j\frac{\pi}{4}}\mathbf{X}_2 + e^{j\frac{\pi}{16}}\mathbf{X}_3 \\ \mathbf{P}_1 e^{j0} &= -\mathbf{X}_0 - e^{j\frac{\pi}{16}}\mathbf{X}_1 + e^{j\frac{\pi}{4}}\mathbf{X}_2 - e^{j\frac{9\pi}{16}}\mathbf{X}_3 \\ \mathbf{P}_2 e^{j0} &= \mathbf{X}_0 - e^{j\frac{9\pi}{16}}\mathbf{X}_1 + e^{j\frac{\pi}{4}}\mathbf{X}_2 - e^{j\frac{\pi}{16}}\mathbf{X}_3 \\ \mathbf{P}_3 e^{j0} &= -\mathbf{X}_0 + e^{j\frac{\pi}{16}}\mathbf{X}_1 + e^{j\frac{\pi}{4}}\mathbf{X}_2 + e^{j\frac{9\pi}{16}}\mathbf{X}_3 \end{aligned}$$

0 00

$$\begin{aligned} \mathbf{S}(t_2) &= \mathbf{P}_0 e^{j0} + \mathbf{P}_1 e^{j0} + \mathbf{P}_2 e^{j0} + \mathbf{P}_3 e^{j0} \\ &= 4e^{j\frac{\pi}{4}}\mathbf{X}_2 = 4\boldsymbol{C}_2\boldsymbol{X}_2 \end{aligned} \quad (17)$$

$$\left(\boldsymbol{c}_k = e^{\frac{j\pi k^2}{16}}\right)$$

Next, calculate $\mathbf{S}(t_1)$.

$$\begin{aligned} \boldsymbol{P}_0 e^{j\frac{\pi}{2}} &= -e^{j\frac{\pi}{2}}\boldsymbol{X}_0 + e^{j\frac{\pi}{16}}\boldsymbol{X}_1 + e^{j\frac{3\pi}{4}}\boldsymbol{X}_2 - e^{j\frac{9\pi}{16}}\boldsymbol{X}_3 \\ \boldsymbol{P}_1 e^{j\pi} &= \boldsymbol{X}_0 + e^{j\frac{\pi}{16}}\boldsymbol{X}_1 - e^{j\frac{\pi}{4}}\boldsymbol{X}_2 + e^{j\frac{9\pi}{16}}\boldsymbol{X}_3 \\ \boldsymbol{P}_2 e^{j\frac{3\pi}{2}} &= e^{j\frac{\pi}{2}}\boldsymbol{X}_0 + e^{j\frac{\pi}{16}}\boldsymbol{X}_1 - e^{j\frac{3\pi}{4}}\boldsymbol{X}_2 - e^{j\frac{9\pi}{16}}\boldsymbol{X}_3 \\ \boldsymbol{P}_3 e^{j2\pi} &= -\boldsymbol{X}_0 + e^{j\frac{\pi}{16}}\boldsymbol{X}_1 + e^{j\frac{\pi}{4}}\boldsymbol{X}_2 + e^{j\frac{9\pi}{16}}\boldsymbol{X}_3 \end{aligned}$$

0 0 0

$$\begin{aligned} \mathbf{S}(t_1) &= \mathbf{P}_0 e^{j\frac{\pi}{2}} + \mathbf{P}_1 e^{j\pi} + \mathbf{P}_2 e^{j\frac{3\pi}{2}} + \mathbf{P}_3 e^{j2\pi} \\ &= 4e^{j\frac{\pi}{16}}\boldsymbol{X}_1 = 4\boldsymbol{C}_1\boldsymbol{X}_1 \end{aligned} \quad (18)$$

$$\left(\boldsymbol{c}_k = e^{\frac{j\pi k^2}{16}}\right)$$

Similarly, about $\mathbf{S}(t_2)$and $\mathbf{S}(t_3)$,

$$\begin{aligned} \mathbf{S}(t_2) &= \mathbf{P}_0 e^{j\pi} + \mathbf{P}_1 e^{j2\pi} + \mathbf{P}_2 e^{j3\pi} + \mathbf{P}_3 e^{j4\pi} \\ &= 4\boldsymbol{X}_0 = 4\boldsymbol{C}_0\boldsymbol{X}_0 \end{aligned} \quad (19)$$

$$\begin{aligned} \mathbf{S}(t_3) &= \mathbf{P}_0 e^{j\frac{3\pi}{2}} + \mathbf{P}_1 e^{j3\pi} + \mathbf{P}_2 e^{j\frac{9\pi}{2}} + \mathbf{P}_3 e^{j6\pi} \\ &= 4e^{j\frac{9\pi}{16}}\boldsymbol{X}_3 = 4\boldsymbol{C}_3\boldsymbol{X}_3 \end{aligned} \quad (20)$$

As a result,

$$\begin{bmatrix} \mathbf{S}(t_0) \\ \mathbf{S}(t_1) \\ \mathbf{S}(t_2) \\ \mathbf{S}(t_3) \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_2\boldsymbol{X}_2 \\ \boldsymbol{C}_1\boldsymbol{X}_1 \\ \boldsymbol{C}_0\boldsymbol{X}_0 \\ \boldsymbol{C}_3\boldsymbol{X}_3 \end{bmatrix} \quad (21)$$

(Ignore the coefficients as they are not important.)

$$|\boldsymbol{S}(t_n)| = \left|\boldsymbol{X}_{\left(\frac{N}{2}-n\right) mod\ N}\right|$$

because of $\left|\boldsymbol{C}_{\left(\frac{N}{2}-n\right) mod\ N}\right| = 1$

Computation is complicated, but many terms cancel clearly each other, resulting in the phase-rotating value after mapping remaining. In sum, we found an astonishing fact, that is the CAZAC precoding makes the amplitudes of M-QAM OFDM signals into the amplitudes of M-QAM single-carrier signals.

3. PROPOSED SYSTEM

As has been noted, CAZAC precoding can make the amplitudes of M-QAM OFDM signals into the amplitudes of M-QAM single-carrier signals. As one application example that can use this characteristic of the waveform of CAZAC-OFDM, we propose inserting a null at a fixed position of the symbol of the time waveform and estimate the symbol timing by detecting the null at the receiving side.

Figure.8 schematically shows the relationship between before and after modulation of CAZAC-OFDM in transmitter side when FFT size is 64.It is arranged in the order of $\mathbf{S}(t_0) = \mathbf{X}(32)\mathbf{C}(32)$,$\mathbf{S}(t_1) = \mathbf{X}(31)\mathbf{C}(31)$, ... ,$\mathbf{S}(t_{32}) = \mathbf{X}(0)\mathbf{C}(0)$,$\mathbf{S}(t_{33}) = \mathbf{X}(63)\mathbf{C}(63)$,…,$\mathbf{S}(t_{63}) = \mathbf{X}(33)\mathbf{C}(33)$ from the beginning of the symbol. This time 2 nulls are embedded at the end of the symbol.

Therefore, the values of $\mathbf{X}(33)$, $\mathbf{X}(34)$ are fixed to 0, and then the transmission waveform of CAZAC-OFDM is generated.Figure9. shows the amplitudes of time domain signals of CAZAC-OFDM (Mapping: QPSK) by the method described above. It is apparent that nulls are inserted firmly at the end of the symbol.

The receiving side judges null and detects the head by the following block configuration(Figure.10). Based on the value obtained by amplitude estimation, the threshold value is set to be half of the distance to the data point closest to 0, using the fact that the transmission waveforms of CAZAC-OFDM are constant amplitudes.

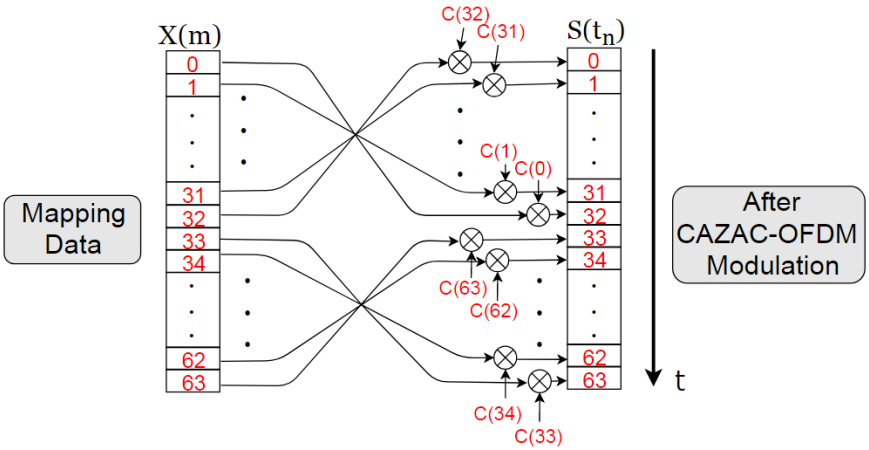


Figure 8. Schematic diagram of the symbol configuration before and after modulation of CAZAC-OFDM (64FFT)

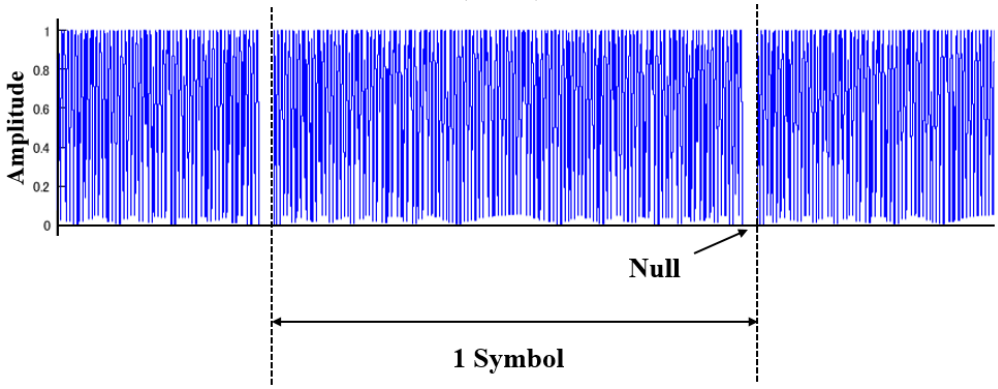


Figure 9. The amplitudes of time domain signals of CAZAC-OFDM embedded Nulls(QPSK)

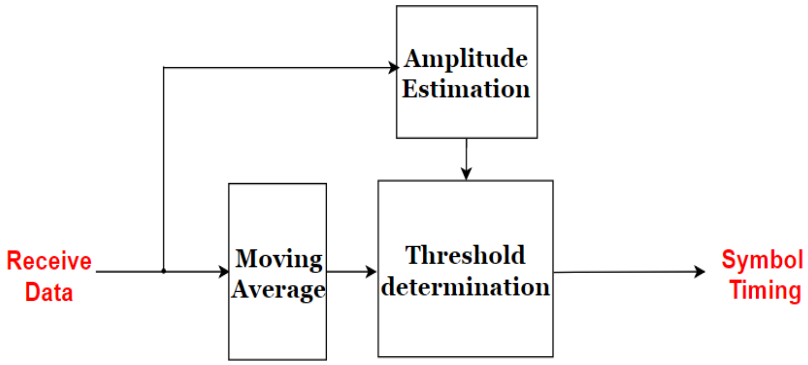


Figure 10.The block configuration of null detection.

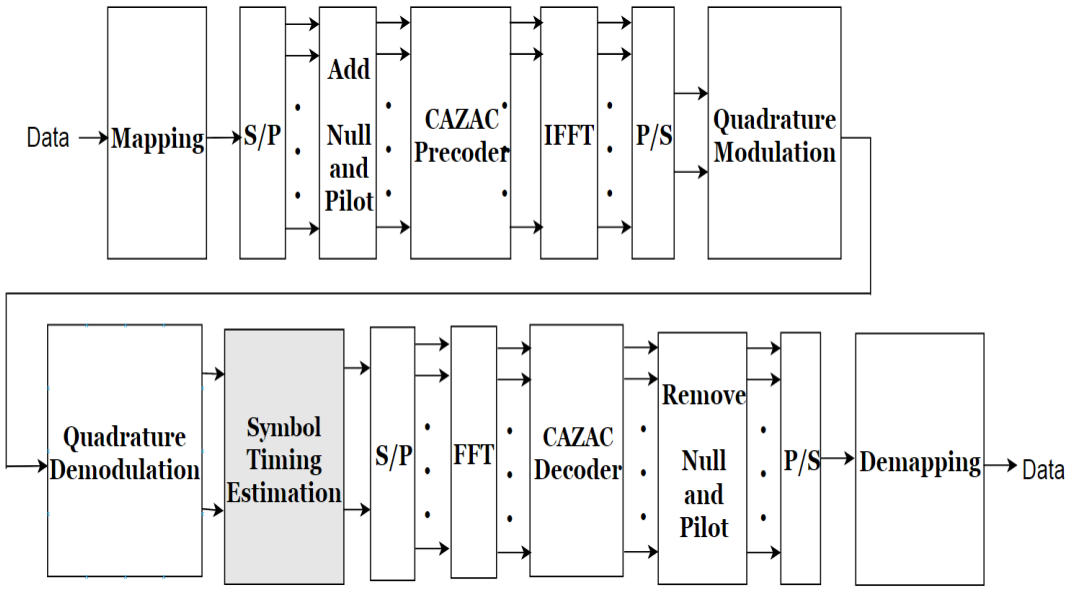


Figure11. The proposed CAZAC-OFDM system

4. SIMULATION RESULTS AND DISCUSSIONS

4.1 Simulation Results

To evaluate the performance of the proposed method, simulation was performed according to the specifications in Table 1. The simulation is performed in MATLAB using communications system toolbox. In addition, head detection using correlation characteristics by guard interval was used as a comparison target. Since the proposed method inserts nulls, it is considered necessary to consider noise. Therefore, among the range where the data itself can demodulate error free in consideration of error correction in this environment (BER $<10^{-3}$), comparison is made at SNR of the worst noise of 15 dB.

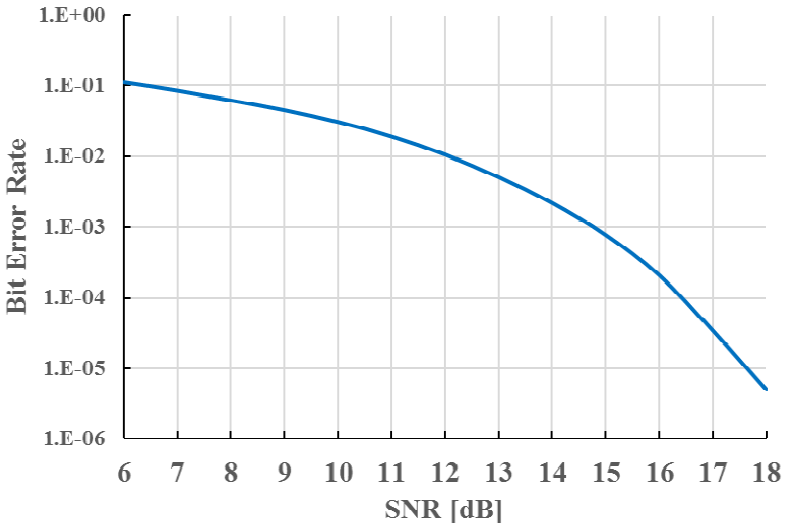


Figure12. BER of simulation work environment.

TABLE 1. SPECIFICATION FOR SIMULATION

	Auto-Correlation (GI)	Proposal
Mapping	16QAM	
Number of Data subcarriers	59	61
Number of Pilot subcarriers	4	1
Number of Null	1	2
FFT size	64	
Number of Guard Interval	16	0
Samples per Symbol	1600	1280
DataRate[Mbps]	43.5125	57.1875
SNR	15 dB	

Figure 13 shows the comparison of detection accuracy with each method. The vertical axis is the timing metric of each method. 0 on the horizontal axis represents the first sample of the symbol. The largest peak appears near 0 in both methods. However, compared with the method using GI, the proposed method shows that a sharp peak appears. From this, it is apparent that the proposed method achieves highly accurate head detection compared to the case of head detection by correlation using GI. In this simulation, 2 samples per symbol were used as synchronization nulls. The throughput at this time is about 30% higher than the throughput when the guard interval of 1/4 symbol length is used.

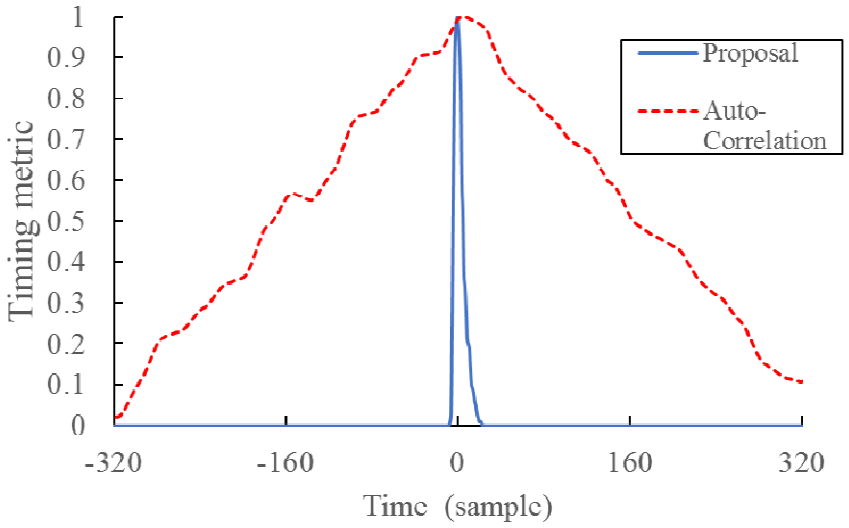


Figure 13. Comparison of detection accuracy.

5. CONCLUSION

In this paper, we have demonstrated that it is possible by using CAZAC-OFDM to make the amplitudes of M-QAM OFDM signals into the amplitudes of M-QAM single-carrier signals. We verified that the OFDM time domain signals can be shaped by using CAZAC precoder. We also verified that, as one application example that can use the characteristic of CAZAC precoder, a new technique of symbol timing estimation provided sufficient performance.

REFERENCES

[1] A. Jovicic, J. Li, T. Richardson" Visible light communication: opportunities, challenges and the path to market", IEEE Communications Magazine, pp.26-32, Dec. 2013

[2] H. Elgala, R. Mesleh, H. Haas:"Indoor broadcasting via white LEDs and OFDM", IEEE Transactions on Consumer Electronics, pp.1127-1134, Aug. 2009

[3] G. Cossu, A. M. Khalid, P. Choudhury, R. Corsini, E. Ciaramella, "3.4 Gbit/s visible optical wireless transmission based on RGB LED", Opt. Exp., pp. B501-B506, Dec. 2012

[4] Z Feng, M Tang, S Fu, L Deng, Q Wu, R Lin, R Wang, "Performance-enhanced direct detection optical OFDM transmission with CAZAC equalization , " IEEE Photonics Technology Letters, pp. 1507-1510,July 2015

[5] I. Baig and V. Jeoti, "PAPR Reduction in OFDM Systems: Zadoff-Chu Matrix Transform Based Pre/Post-Coding Techniques," in Proc.of the 2nd International Conference on Computational Intelligence, Communication Systems and Networks, pp. 373-377, July 2010.

[6] R. Ishioka, T. Kimura and M. Muraguchi, "A Proposal for a New OFDM Wireless System using a CAZAC Precoding Scheme,"Proc. AICT 2017, pp. 47-51, June 2017.

[7] Kazuki Miyazawa, Tomotaka Kimura, Masahiro Muraguchi, "Proposal of visible light OFDM system with CAZAC equalization," 23rd Asia-Pacific Conference on Communications (APCC), pp.491-496, Dec. 2017.

[8] T.M. Schmidl, D.C. Cox:" Robust frequency and timing synchronization for OFDM", IEEE Transactions on Communications, pp.1613-1621, Dec. 1997

[9] B.Park, H.Cheon, C. Kang, D. Hong:"A novel timing estimation method for OFDM systems", IEEE Communications Latters, pp. 239-241, May 2003

INTENTIONAL BLANK

IOT-BASED HOME APPLIANCE SYSTEM (SMART FAN)

[1] Mehran Ektesabi, Saman A. Gorji, Amir Moradi,
[2] Suchart Yammen, V. Mahesh K. Reddy,
[3] Sureerat Tang

[1]Swinburne University of Technology, Victoria, Australia
[2]Naresuan University, Thailand
[3]TJ Supply Limited Partnership, Thailand

ABSTRACT

Smart home appliances such as smart fridge, smart lighting, and smart air conditioner are getting popular for home end users. Smart fans as one of those smart devices are a part of a smart home that can be assumed as a factor of comfort, which may also reduce the electricity cost due to its high efficiency.

Hence, this project aims to develop an alternative smart fan tackled from a comfort and cost perspectives. This project is done using as minimum budget as possible by using a combination of the already-available parts of the market.

It is expected to develop a prototype of a cheap smart fan, which in turn becomes the starting point to allow further development of other smart home appliances

KEYWORDS

Internet of Things, Smart Fan, Motor Control, Cost Analysis, Cloud Infrastructure

1. INTRODUCTION

Automation has become one of the key interests in the modern-day technology. Everybody tends to use automated devices in his or her daily activities due to several reasons ranging from safety to ease of handling. Initially, the automation systems were limited to industries as it required significant investments, but with the development of the technology, automation has become available to everyone. Home automation systems are very popular in the world these days. Many types of research have been carried out into this area, and some commercial products are also available, but the room for development is still very high, and the technology advances daily. Unlike the normal system, the automated home will have additional benefits. As the electric system is automated, it will save electricity by switching off the appliances when they are not required. On the other hand, as the appliances can be remotely controlled, the user can decide what to do even he is not at home. The systems can be developed to a level where the security of the house can be improved. One of the famous home automation solutions is to use a dedicated system connected to an application installed on the user's smartphone. The combination of these two hardware and software allows greater control over home automation and put it all in the user's hand.

Natarajan Meghanathan et al. (Eds) : AIAA,WiMNeT, DPPR - 2018
pp. 37–46, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.81604

Smart Technology has been chosen by many researchers and worked on different methodologies to achieve their objectives. As an instance a team worked on Smart fan robot which detects and tracks a person automatically using thermal camera and RGB-D camera [1].Another research study was done on human tracking system by using ultrasonic distance sensors to detect one person only up to 3 meters [2]. PWM technique is also used to control the fan speed by changing the duty cycle according to room temperature [3]. Also a group have developed a smart fan which operates according to the presence of human, ambient temperature and the position of humans [4]. Another project is done on controlling home appliances using smart phone with Wi-Fi as communication protocol and Raspberry Pi as a server system [5]. Another research team have conducted different tests and concluded that the fan speed can controlled by room temperature using PT-100 sensor [6]. Similar work was done on automatic fan speed control using PIC microcontroller. The authors used a special feature called Buzzer which generates alarm if the temperature becomes overheat [7]. Another group have designed a room temperature and humidity controller by using two fuzzy logic controllers. Their model is operated in between 8 and 44 ^{0}C rage of temperature and they proved that their model works properly between 18 and 26 ^{0}C temperature range in order to maintain the user comfort level [8]. The power electronics system can aid to implement more efficient systems [9]-[10]. In this area, a study has introduced a new method of speed control of induction motor by connecting power electronic circuit. This shows the increase in efficiency of the fan motor at low speeds and also starting torque is high at any speed value [11]. As an extension to these works done so far, in this project we interfaced Internet of Things (IoT) to the fan speed control by conducting the following techniques: IoT based motor control, Zero crossing Detection, Phase Angle Control (Triac Control) and Closed Loop Control to achieve the automatic speed control of the fan in an efficient way which becomes more user friendly to the human.

2. SYSTEM REQUIREMENTS

Automation has become one of the key areas which world focus at present. People tend to modify things which are often used so that they can be controlled by themselves or remotely by human intervention. With the invention of smartphones everything on earth has become possible with few clicks. In that type of an environment, it is important to have a user-friendly system where the devices in a house can be remotely controlled using a mobile phone. Through this project, we expect to make a smart fan which can be controlled via Bluetooth using an android application. Day to day normal room fans does not take the room temperature in to account we have to manually control the speed of the fan according to the temperature. We try to automate this system where the speed of the motor of the fan will change according to the temperature of the room. Table 1 shows the problem/need statement of this project.

Table 1. The project need statement

Element	Description
The problem of existing conventional Fans	In case of ceiling fan, table fans or pedal fans etc., the speed can be varied at step wise only three steps. The user needs to control it manually which makes difficult in some cases. When the user needs to increase or decrease the speed based on his comfort level, he or she operates it manually every time. In addition, the power consumption is high, and air flow is less in conventional fans.
Affects	Some users like aged persons, physically handicapped people mainly feel problematic to control the fan every time manually when they need change in air flow.
And results in	User feels discomfort with this type of drawbacks; therefore, the manufacturing company loses its stakeholders.
Benefits of a solution	The proposed control method solves these problems where user can control fan remotely and also improves the customer comfort levels automatically.

3. ARCHITECTURE DESIGN

For an IoT device to work, it required an architecture of hardware and software that supports the operation of IoT devices. This has been elaborated in the following subsections:

The supporting architecture could be categorised into many forms, depending on which part the focus lies. One of the categories is based on whether it is a hardware or software. Hardware architecture including but not limited to the gateway, cloud infrastructures, and even connectivity to the cloud interfaces. Software architecture including device management, the operating system of any hardware infrastructures, and also including communication protocol. Although there are many components which exists in the supporting architecture, we will focus more on the elements which are the core of IoT applications.

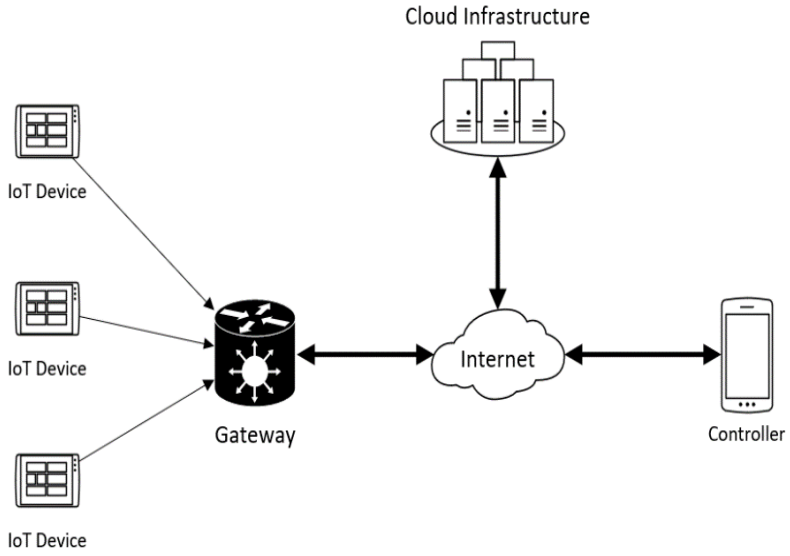


Figure 1. Generic Connection Architecture of an IoT Application

For small-scale IoT applications – such as temperature aware smart fan – does not need many supporting platforms. Some of the supporting platforms is not essential to the operation, and excluding those things could improve the development time of the system. Elements such as device management, provisioning, access controller, and resource allocation are useful, but its presence is not critical to a small-scale IoT application.

Small-scale IoT applications usually have a connection architecture as one placed in Figure 1. For this type of architecture, there are four essential elements: IoT Device, Gateway, Cloud Infrastructure, and Controller.

1. IoT Device

This device acts as the main element of the IoT applications. IoT device is a device which connects the whole IoT application to the physical world. Some examples of this connection are a temperature sensor, motor controller, building alarm, or even a webcam.

2. Gateway

The gateway acts as a middle layer service between IoT devices, with the internet or another type of communication carrier to the cloud infrastructure. This device is usually also capable of communication processing, and also device management. Depending on the gateway design, this device could reach all layers in OSI model. So, the function of the gateway is various; it depends on the applications and the system's requirements.

3. Cloud Infrastructure

This element stores and process the data for/from the IoT devices. Depending on the application, this element could function as a database, or even as a smart decision server.

4. Controller

This element acts as the user interface for the system. Users could control and manage the behaviour of the applications.

4. EXPERIMENTAL RESULTS

4.1. Motor Controller Results

The motor was tested by TRIode for Alternating Current (TRIAC). Because Triac controller circuit runs on 240 V, there are two steps for testing this Triac-based controller. The first one is by testing it on 20 Vpp, and the latter part is testing on 240 RMS live voltage.

The first part used a signal generator as the voltage source, and an oscilloscope is used to read the voltage level at a different point in the circuit.

Next is to test the output of the triac. For this test, the firing is set to be constant at 90 degrees, and the voltage across triac terminal is observed on the oscilloscope. In Figure 2.a- Figure 2.b, there are three signals: blue for zero-crossing, purple for triac firing control, and green for the terminal-to-terminal voltage. As can be seen in the graph, when the purple signal is set to high, the terminal voltage (green) drops to a constant. This result concluded that the triac is working.

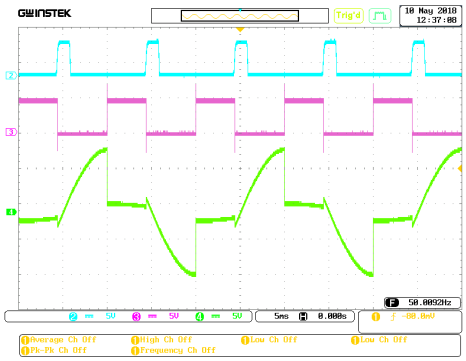


Figure 2.a Triac Functional Testing (90 degrees)

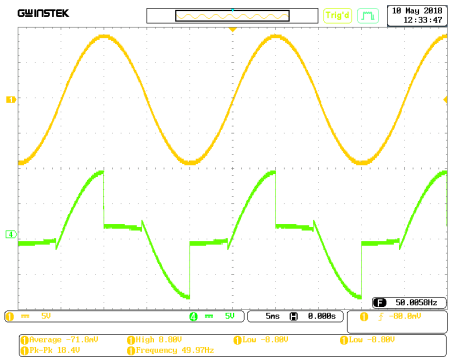


Figure 2.b Triac Terminal vs Source Voltage

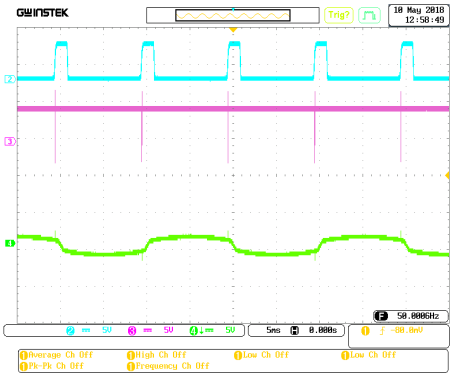


Figure 2.c Test with Firing Angle at 0 degree

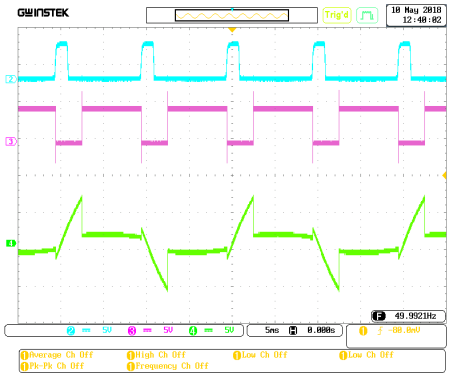


Figure 2.d Test with Firing Angle at 45 degrees

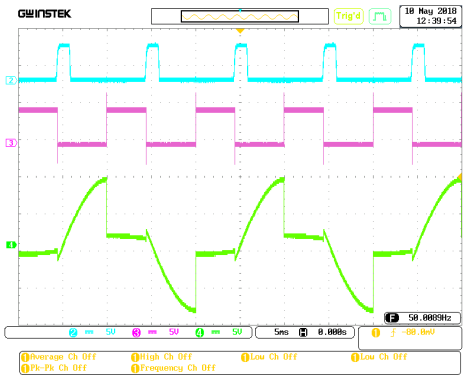


Figure 2.e Test with Firing Angle at 135 degrees

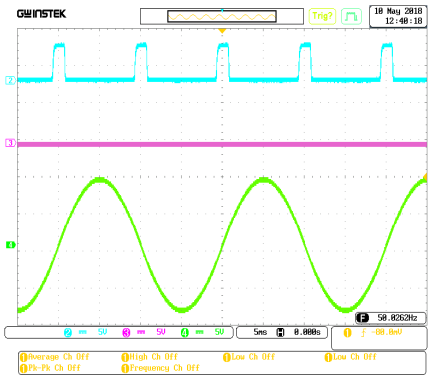


Figure 2.f Test with Firing Angle at 180 degrees

Figure 2. Motor control experiments

In Figure 2. c, Figure 2.d, Figure 2.e and Figure 2.f, the terminal voltage switched to a constant voltage at angle value close to 0°,45°,135°,and 180° respectively. Finally, at 180°, the terminal voltage would be expected to not open at any angle which is obtained from Figure 2.f.

4.2. Android Controller

Testing of the Android controller is done in two parts: modular testing and full-system integration testing, which will be done in the next section.

For modular testing, the application was tested to interact with a single device at a time: either to the sensor or the motor. Figure 3.a contains the temperature reading from the sensor either from the sensor itself or from the manual polling via device list options in Figure 3.b. There is also a status report which is based on the last update timestamp from the sensor (Figure 3.a, second card).

Figure 3.a Android Controller - Main View

Figure 3.b Android Controller - Device List

Figure 3. App/Sensor experiments

4.3. System integration

This section explains the full-system integration test. The system is set up using an additional heater fan as the heating element to control the temperature reading of the sensor. The test begins with the heating fan heats up the temperature sensor to a temperature high enough so that the fan could cool down the temperature (i.e. much higher than the ambient temperature). Figure 4.a, shows the initial temperature reading after heating up the sensor. Then, the fan is turned on at full speed to cool down the sensor. After cooling down several degrees Celsius, the fan speed is slightly reduced as shown on Figure 4.b, because the difference in temperature and set point is now reduced

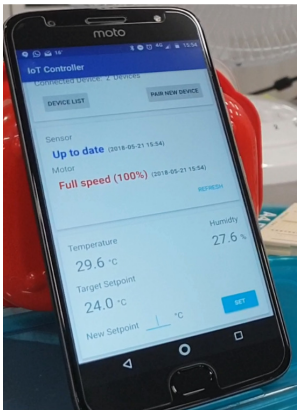


Figure 4.a Test – Initial State

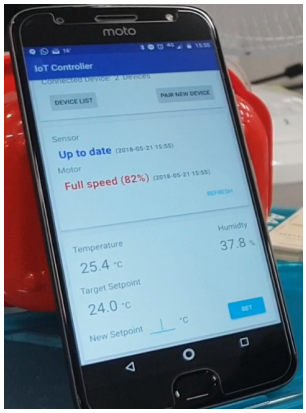


Figure 4.b Test – Fan Full Speed

Figure 4. System tests

5. MARKET STUDY

Entrepreneurs and investors look for business opportunities with potential growing market [12]. Internet of Things (IoT) has a great potential to attract new and existing businesses. Where there is a potential market, analysing the profitability of a new product is a crucial steps in establishing and expanding a business. To achieve the expected rate of return an accurate product cost estimation is the key. It provides managers and investors with necessary information required for strategic decisions and achievement of business objectives. Mainly, in today's competitive environment, it is critical to have reliable product cost estimations for pricing decisions.

5.1 Cost analysis

To estimate our product cost, we use the total absorption costing method. Accordingly, the cost is divided to direct and indirect costs. Specifically, our product cost estimation consists of direct material, direct labour and overhead. The required direct materials and their costs are estimated as follows:

Table 2. breakdown of the component costs

Description	Qty	Unit Cost (USD)*	Unit Cost (AUD)	Total Cost (AUD)
Conventional Fan	1		14	14
Arduino Uno - R3 SMD	1	29.95		38.935
Temp. Sensor DHT-22	1		9.95	9.95
SparkFun Bluetooth Modem - BlueSMiRF Silver	1	24.95		32.435
Breadboard Jumper Wire Bundle	1		4.95	4.95
Female Headers	2	1.5		3.9
Break Away Headers - Straight	2	1.5		3.9
Subtotal				$108.07

*USD Rate 1 USD = 1.3 AUD

In our preliminary analysis, direct labour and overhead costs are estimated based on this assumption that the product will be produced in Australia (the product cost will vary if the location changes. The direct labour cost is estimated based on one hour assembly time at the rate of $20. In addition, the estimated overhead costs are as follows for 1000 to 2000 units per year:

Table 3. breakdown of the manufacturing costs

Rent	20,000
Salaries	50,000
Equipment depreciation	10,000
Utilities	10,000
Other costs	20,000
Total fixed overhead	$110,000

In early stages of introducing a product, penetration pricing can be used as a strategy to attract customers through word of mouth [13]. Penetration pricing is setting a low price in the beginning of marketing campaign to gain a high market share. To attract more customers, for the first year we assume that we do not expect any profit and want to determine the selling price at a level that we can only cover all costs (both fixed and variable). To do so we use the break-even analysis technique. Break-even point is the number of products should be sold to cover all costs while making no profit. Break-even point analysis is a useful method to not only profit projection but also for pricing decision and expense control [14]. To calculate the break-even point the following items are required: fixed cost, selling price and variable cost.

Accordingly, our suggested selling price is almost $185 and break –even point is 1932 units.

Table 4. price calculations

Direct material (variable cost)	$108.07
Direct labour (variable cost)	$20
Overhead per unit	$57
Unit cost	$185

This cost estimation shows that if we produce and sell 1,932 units at the price of $185, all costs will be covered.

As this pricing decision is made only for introducing the product to the market for the first year, different strategies are required to make profit in following years. To determine the alternatives we first discuss the value of our product from the customer point of view.

5.2. Product values

The Internet of Things can transform our various aspects of lives and enhance both efficiency of devices and comfort of individuals [15]. To analyse the impact of efficiency improvement of our product, the financial benefits that a customer will receive from using the product over a year are estimated. To do so, we first calculate the electricity cost of running a traditional 100w fan for the whole summer (90 days, 10 hours per day) for a household in Australia.

Table 5. Energy costs

Hours used	900
Energy consumption(Kw)	90kW
Price (kWh)	0.35
Electricity cost	$31.5

Our smart product might be able to reduce this electricity cost by 50%. Comparing our estimated product price and the potential cost saving shows a significant gap. Specifically, more than 10 years electricity saving is required to justify the price. Thus, it can be argued that the electricity cost saving cannot attract customers and other values such as customers' comfort should be the focus which can affect our pricing model as well.

6. CONCLUSIONS

The summary of results and conclusion is basically divided into two sections that are as follows.

6.1. Automation Section Conclusion

A novel IoT based Fan speed control system has been designed and tested using IoT concepts. A computer program was developed to control the speed of the fan by using smart phone and is successfully implemented. The fan motor is operated and tested at different speed input value from the smart phone and the experimental results are noted. These results show that the user can switch on the fan, regulate its speed by using the smart phone. The user can operate this fan even he/she is in remote location. By using temperature and humidity sensors the speed control of fan becomes very easy so that it automatically changes its speed when the temperature is changed according to the user's comfort zone. All the values related to temperature, humidity and speed of the motor are also stored in the cloud storage, and the software can display these data for operators to check. The main advantage of this method of control is that the user can control the

speed of the fan by giving the set point value via the smart phone. Therefore, the fan motor can be efficiently controlled by using the Wi-Fi directly from anywhere and anytime in this world with proper internet connection.

6.2. Business Section Conclusion

The preliminary cost analysis of our product shows that producing the product in Australia is not viable and further analysis is required to estimate the cost of production in Asian countries. Moreover, the monetary value of electricity cost saving of our product is not significant as the electricity consumption of a traditional fan is relatively low. Thus, the focus needs to be on the comfort aspects that our product can provide as well as adopting the technology to high-wattage devices. Specifically, the product can be used for patients, elderly or disabled people who cannot adjust the temperature manually. For example, it might not be possible for a person with disability to adjust the fan manually but using our product the adjustment will incur automatically and there is no need for human intervention. Finally, we suggest the use of our developed technology in this project for other electrical devices such as heaters and coolers with higher electricity consumption to provide both efficiency and comfort.

ACKNOWLEDGEMENTS

This research was supported by Swinburne University of Technology in collaboration with TJ Supply Limited Partnership under Seed Grant 35818.

REFERENCES

[1] S. Shimamura, K.Matsumoto, N.Maeda, T.Kodera, W. Nakagawa, Y. Shinozuka, M.Sugimoto and H.Saito, " Smart Fan: Self-contained Mobile Robot that performs Human detection and tracking using thermal Camera," International Conference on Artificial Reality and Telexistence Eurographics Symposium on Virtual Environments, 2014.

[2] Tajrin Ishrat, Mohammad Anisur Rahaman and Arif Ahammad, "Smart Fan for Human Tracking," the 9th International Forum on Strategic Technology (IFOST), October 21-23, 2014, Cox's Bazar, Bangladesh.

[3] Vaibhav Bhatia and Gavish Bhatia, "Room temperature based Fan Speed Control System using Pulse Width Modulation Technique," International Journal of Computer Applications (0975-8887) Volume 81-No5, November 2013.

[4] Md. Mozasser Rahman, Mohd Fahrul Radzi Bin Zakaria and Shahrul Na'im Sidek, "Sensory and Control System for Smart Fan," International Journal of Control, Automation and Systems, Vol.4 No.3, July2015, ISSN 2165-8277 (Print), ISSN 2165-8285 (Online).

[5] Pavithra. D and Ranjith Balakrishnan, "IoT based Monitoring and Control System for Hom Automation," Proceedings of 2015 Global conference on Communication Technologies (GCCT 2015).

[6] M.A.A. Mashud, Dilruba Yasmin, M.A. Razzaque and M.H. Uddin, "Automatic Room Temperature Controlled Fan Speed Controller using PT-100," International Journal of Scientific & Engineering Research, Volume 6, Issue8, August-2015, ISSn 2229-5518.

[7] Zairi Ismael Rizman, Kim Ho Yeap, Nuraiza Ismail, Norizan Mohamad and Nur hafizah Rabi'ah Husin, "Design of Automatic Temperature Control System for Smart Fan using PIC," international Journal of Science and Research (IJSR), India Online ISSN:2319-7064.

[8] Tarun Kumar Das and Yudhajit Das, “Design of a Room Temperature and Humidity Controller using Fuzzy Logic,” American Journal of Engineering Research (AJER) e-ISSn:2320-0847 p-ISSN:2320-0936, Volume-02, Issue-11, pp-86-97.

[9] S. A. Gorji, M. Ektesabi and J. Zheng, "Isolated switched-boost push–pull DC–DC converter for step-up applications," in Electronics Letters, vol. 53, no. 3, pp. 177-179, 2 2 2017.

[10] S. Asghari and E. Fallah, "A new approach for efficiency optimizing of single-phase induction motors," 2012 3rd Power Electronics and Drive Systems Technology (PEDSTC), Tehran, 2012, pp. 500-505.

[11] Hamid M.B. Metwally, “New method for speed control of single phase induction motor with improved motor performance,” Energy Conversion and Management 42 (2001) 941-950.

[12] Sahlman, W.A., 2008. How to write a great business plan (No. E70 90). Harvard Business School Press.

[13] Yenipazarli, A., 2015. A road map to new product success: Warranty, advertisement and price. Annals of Operations Research, 226(1), pp.669-694.

[14] Dean, J., 1948. Cost structures of enterprises and break-even charts. The American Economic Review, pp.153-164.

[15] Stojkoska, B.L.R. and Trivodaliev, K.V., 2017. A review of Internet of Things for smart home: Challenges and solutions. Journal of Cleaner Production, 140, pp.1454-1464.

A PAPR REDUCTION TECHNIQUE IN OFDM SYSTEMS WITH A LARGE NUMBER OF SUBCARRIERS

Yasuhiro Shimazu[1], Yushi Shirato[2] and Masahiro Muraguchi[3]

[1,3]Department of Electrical Engineering, Tokyo University of Science
6-3-1 Niijuku, Katsushika-ku, Tokyo, 125-0051, Japan
[2]NTT, Japan

ABSTRACT

A major drawback of orthogonal frequency division multiplexing (OFDM) signals is extremely high peak-to-average power ratio (PAPR). Signals with high PAPR lead to a lowering of the energy efficiency of power amplifiers and the shortened operation time causes a serious problem in battery-powered wireless terminals. We have found the CAZAC precoding makes the PAPR of M-array quadrature amplitude modulation (M-QAM) OFDM signals into the PAPR of M-QAM single-carrier signals. Therefore, it can dramatically improve the PAPR of OFDM signals. However, to satisfy the 3GPP-LTE specification of frequency spectrum, severe bandpass filtering of CAZAC-OFDM signal lead to unacceptable regrowth of the PAPR. The paper provides available control procedure for PAPR and spectrum managements. It is confirmed that the CAZAC-OFDM signal controlled by our procedure maintains enough low PAPR and can provide comparable spectral specifications in the downlink channel of 3GPP-LTE standard.

KEYWORDS

OFDM, CAZAC, PAPR, spectral specifications

1. INTRODUCTION

Orthogonal frequency division multiplex (OFDM) systems that attain high speeds and high capacity have recently been attracting attention in wireless applications, e.g., wireless local area networks (WLANs), third generation partnership project long-term evolution (3GPP LTE), and the digital video broadcasting-terrestrial (DVB-T) standard [1], [2]. The main drawback of OFDM signal is its high peak-to-average power ratio (PAPR), which decreases the efficiency of power amplifiers (PAs) and increases transmitter power consumption [3]. Therefore, many techniques have been proposed to reduce the PAPR [4]. Well-known techniques are clipping-and-filtering, partial transmit sequences (PTSs), and selected mapping (SLM). The clipping-and-filtering limits the peak amplitude of the transmission signal. Since the technique, however, cut a portion of signal amplitude with important information, as a result, it causes non-linear distortion and degrades the BER performance. The PTS partitions input data into disjoint sub-blocks. Moreover, each sub-block is weighted by a phase factor. This technique chooses the phase factor to minimize the PAPR of combined signals. The SLM generates multiple candidate data blocks. All data blocks represent the same information. Although the PTS and the SLM can be expected to create a certain reduction in PAPR, both techniques need side information in the receiver, which decreases spectral efficiency. The most practical solution to improving PAPR is to

Natarajan Meghanathan et al. (Eds) : AIAA, WiMNeT, DPPR - 2018
pp. 47–61, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.81605

introduce a single carrier scheme, i.e., single carrier frequency division multiplexing access (SC-FDMA). The 3GPP LTE system adopts SC-FDMA for uplink multiple access systems [5]. However, SC-FDMA is not the best solution from the viewpoint of high spectrum utilization.

As the other option of PAPR reduction techniques, the constant amplitude zero auto-correlation (CAZAC) precoding had been proposed [6]. Afterwards we had found out that the CAZAC precoding (CAZAC-OFDM) made the PAPR of M-array quadrature amplitude modulation (M-QAM) OFDM signals into the PAPR of M-QAM single-carrier signals [7]. Moreover, we had demonstrated the theoretical verification of the CAZAC-OFDM [8].

To put the CAZAC-OFDM into practical use, we are facing new complex problems. Although the CAZAC precoding dramatically improves the PAPR of OFDM signals, in order to satisfy the 3GPP-LTE specification of frequency spectrum, severe bandpass filtering of CAZAC-OFDM signal lead to unacceptable regrowth of the PAPR.

The paper provides available control procedure for PAPR and spectrum managements for the CAZAC-OFDM system. One simplest approach of improving the PAPR is to clip the regrowth amplitude of the signal to a fixed level. Here, as the regrowth amplitude has not important information, the BER performance does not been degraded. An accurate clipping boundary is easily defined, because we can manage the amplitude of the time-domain signal of CAZAC-OFDM like a single-carrier signal. As the correct demodulation of CAZAC-OFDM signals requires only values of the original signal points, we can improve the PAPR without any degradation of the BER and little degradation of the spectrum by the clipping. The CAZAC-OFDM scheme provides high-efficiency operation of power amplifiers, i.e., the Power-Added-Efficiency (PAE) of 38%, the power consumption of the PAs becomes about one-third of those of conventional OFDM scheme, and it is comparable value of single carrier signal scheme, i.e., the SC-FDM scheme. Furthermore, although the PAPR of conventional OFDM scheme strongly deteriorates as increasing the number of subcarriers, the PAPR of the CAZAC-OFDM scheme does not depend on the number of subcarriers. We already confirmed the usefulness of CAZAC-OFDM for the case of 64 subcarriers, and in this paper, we examined CAZAC-OFDM using 1024 subcarriers closer to the LTE specification [9]. Therefore, we believe the CAZAC-OFDM scheme can use for uplink channels of future wireless mobile communication systems.

2. CAZAC-OFDM SYSTEM

2.1. OFDM System

In OFDM system, the frequency domain symbol $\boldsymbol{X} = [X_0, X_1, \dots, X_{N-1}]^T$ is modulated by N size inverse Fast Fourier Transform (IFFT). The discrete-time OFDM signal with N subcarriers are represented as

$$x_n = \sum_{k=0}^{N-1} X_k e^{j2\pi kn/N}, \tag{1}$$

where $j = \sqrt{-1}$ and n is discrete time index. On the other hand, receiver acquires frequency domain symbol $\boldsymbol{Y}$ by applying FFT to received signal $\boldsymbol{y}$.

$$Y_k = \sum_{n=0}^{N-1} y_n e^{-j2\pi kn/N} \tag{2}$$

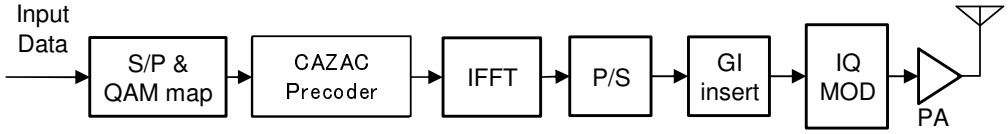
$$= \sum_{n=0}^{N-1} (x_n + Noise) e^{-j2\pi kn/N}.$$

The PAPR of the OFDM signal eq. (1) can be expressed as

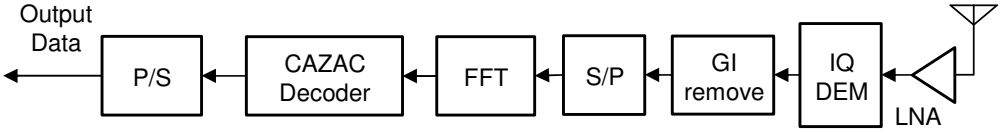
$$PAPR = \frac{\max\limits_{0 \le n \le N-1} |x_n|^2}{\mathrm{E}[|x_n|^2]}, \tag{3}$$

where E[·] is expectation operator. PAPR represents amplitude fluctuation of each symbol. To improve the accuracy of PAPR, the OFDM signal x_k is converted to L-times oversampled time domain signal.

As shown from eq. (2), the OFDM signal is composed of a plurality of subcarrier signals, which causes an increase in amplitude fluctuation. A high PAPR signal increases the Input Back Off (IBO) at the power amplifier to amplify the transmit signal without distortion. In general, increasing in IBO causes decreasing the efficiency of PA [3].



(a) Transmitter



(b) Receiver

Figure 1.CAZAC-OFDM system

2.2. CAZAC Precoding Technique

Figure 1 shows a transmitter and receiver block diagrams for a CAZAC-OFDM system. The CAZAC sequence is an orthogonal sequence and the autocorrelation function is a delta function. In addition, the CAZAC sequence has a characteristic that the amplitude exhibits a constant value in both time and frequency domain. Zadoff-Chu (ZC) sequences is one of the CAZAC sequence. The ZC sequence $C_k$of length L is defined as

$$C_k = \begin{cases} exp\left(\dfrac{j\pi k(k-1)r}{L}\right) (L \text{ is even}) \\ exp\left(\dfrac{j\pi (k-1)^2 r}{L}\right) (L \text{ is odd}) \end{cases} \tag{4}$$

Where L is the length of the CAZAC sequence and r is the sequence number, $k = 1, 2, \cdots L$. L is a natural number, and r is a prime integer with respect to L. CAZAC sequences are generated by cyclic shift of the original CAZAC sequence. The periodic cross-correlation function ρ is defined as

$$\rho(m) = \sum_{n=1}^{L-1} c_n c^*_{(c-m) \bmod L}$$

$$= \begin{cases} L\ (m = 0) \\ 0\ (m \neq 0) \end{cases} \tag{5}$$

Where m is integer variables.

CAZAC precoding uses a square matrix M generated from the equation in the case where L in equation (4) is an even number. The matrix equation is defined as

$$M = \frac{1}{\sqrt{N}} \begin{bmatrix} c_1 & c_2 & \cdots & c_N \\ c_{N+1} & c_{N+2} & \cdots & c_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ c_{(N-1)N+1} & \cdots & \cdots & c_{N^2} \end{bmatrix} \tag{6}$$

Where matrix M is the rearrangement of in equation (4) in the row direction, and N is the number of subcarriers and $L = N^2$, $r = 1$.N is an even number, so L is also an even number.

Multiply the signal before IFFT by the matrix calculate the product of the transposed the frequency domain symbol $X = [X_1, X_2, \cdots, X_N]$ and the matrix M, and create a CAZAC equalized signal X'. X' is represented as

$$X' = M \cdot X^T$$

$$= \frac{1}{\sqrt{N}} \begin{bmatrix} c_1 & \cdots & c_N \\ c_{N+1} & \cdots & c_{2N} \\ \vdots & \ddots & \vdots \\ c_{(N-1)N+1} & \cdots & c_{N^2} \end{bmatrix} \cdot \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} \tag{7}$$

An IFFT operation is performed on this signal. The OFDM time domain signal $x_n$of the sample is represented as

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X'_k e^{j2\pi kn/N}$$

$$= \sum_{k=0}^{N-1} \left\{ \sum_{m=0}^{N-1} e^{j\pi(m+kN)^2/L} X_m \right\} e^{j2\pi kn/N}$$

$$= \sum_{m=0}^{N-1} e^{j\pi m^2/N^2} X_m \left\{ \sum_{k=0}^{N-1} e^{j2\pi k(m+n)/N} e^{j\pi k^2} \right\} \tag{8}$$

Where k is an integer not less than 0, so the following equation is developed.

$$exp(j\pi k^2) = \begin{cases} 1 & (k : \text{even}) \\ -1 & (k : \text{odd}) \end{cases} \tag{9}$$

To lead the equation (10) from the equation (9).

$$\exp(j\pi k^2) = (-1)^k \tag{10}$$

Substituting the equation (10) into the equation (8) lead to the equation (11).

$$x_n = \sum_{m=0}^{N-1} e^{j\pi m^2/N^2} X_m \left\{ \sum_{k=0}^{N-1} \{-e^{j2\pi(m+n)/N}\}^k \right\} \tag{11}$$

The inside of {} in equation (11) is the sum of the geometric progression. Therefore, equation (12) is derived.

$$\sum_{k=0}^{N-1} \{-e^{j2\pi(m+n)/N}\}^k = \begin{cases} N & (-e^{j2\pi(m+n)/N} = 1) \\ 0 & (-e^{j2\pi(m+n)/N} \neq 1) \end{cases} \tag{12}$$

When $2\,\pi\,(m+n)\,/\,N = 1$, $2(m+n)/N$ is an integer and odd number. Also, n and $m$are$0 \leq n \leq N-1, 0 \leq m \leq N-1$.These relationships are satisfying the relationship of equation (13).

$$m = \frac{\mathrm{N}}{2} - n\ (mod\ N) \tag{13}$$

Equation (14) is derived from equations (11), (12) and (13).

$$x_n = e^{j\pi\{\mathrm{N}/2 - n\ (mod\ N)\}^2/N^2} X_{\frac{N}{2} - n(mod\ N)}$$

$$= c_{\frac{N}{2} - n(mod\ N)} X_{\frac{N}{2} - n(mod\ N)} \tag{14}$$

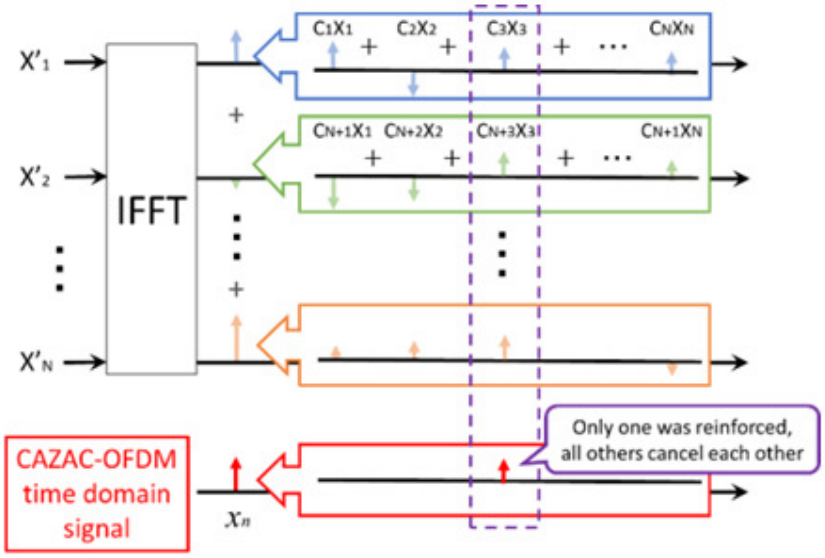


Figure 2.The time domain signal of CAZAC-OFDM[8]

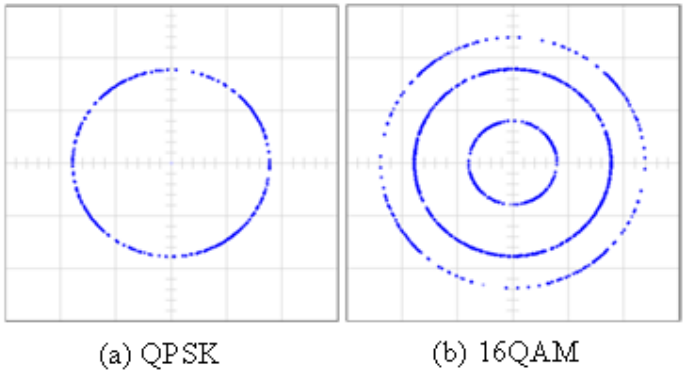


Figure 3.The time domain signal points of
the CAZAC-OFDM Signals on the complex phase plane.

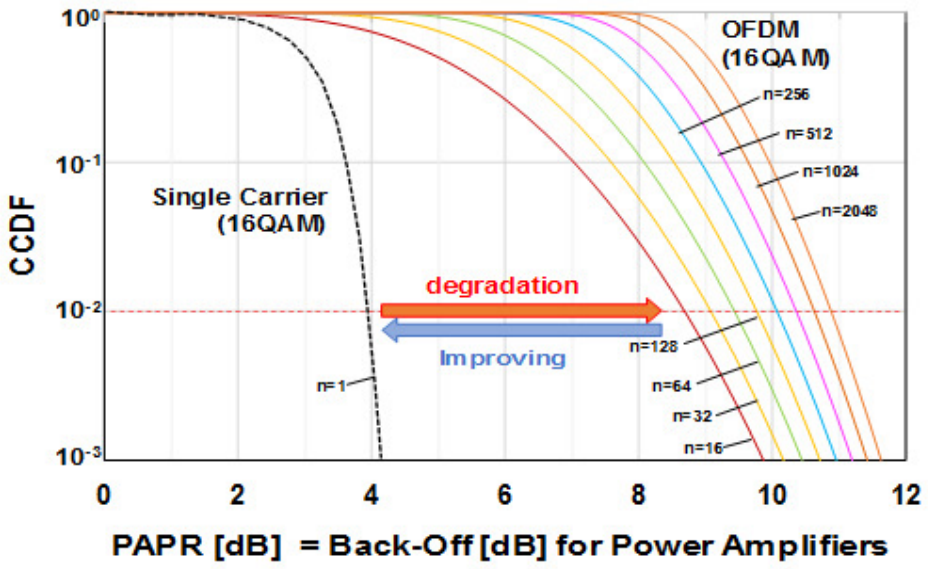


Figure 4. Demerit of increasing number of subcarriers in OFDM

Equation (14) shows that CAZAC precoding converts the PAPR of the OFDM time domain signal into the PAPR of a single-carrier signal. Figure 2 shows the image diagram of the CAZAC-OFDM time domain signal. Among the components of the coefficient X, only one was reinforced, all others cancel each other.

Figure 3 (a) and (b) show the time domain signal points of CAZAC-OFDM, which shows the IQ-MOD input in Fig. 1. The discrete-time signal points of QPSK CAZAC-OFDM line up on the unit circle orbit as shown in Fig. 3 (a), and the discrete-time signal points of 16QAM CAZAC-OFDM signal are on the three concentric circle orbits as shown in Fig. 3 (b).

Figure 4 shows demerit of increasing number of subcarriers in OFDM. In the conventional OFDM, PAPR also deteriorates as the number of subcarriers increases, but CAZAC-OFDM can improve PAPR as much as single carrier without depending on the number of subcarriers. This characteristic is very suitable for the recent trend of increasing the communication speed by increasing the number of subcarriers.

3. SPECTRAL CONTROL

3.1. Procedure of Spectral Control

To make the discrete-time OFDM signal with the proper pass band characteristics, we set the subcarrier allocation as shown in Fig.5. This process is equivalent to severe filtering. That is, while spectral control can be performed, transient response occurs between signal points. However, due to oversampling, this transient response part appears as an interpolation point between the original signal points in the time domain. Here, we set the FFT size of CAZAC-OFDM to 1024. Therefore, four times interpolated (oversampled) FFT size becomes to 4096 and three interpolation points are inserted between the original signal points.

The CAZAC-OFDM signal at the IQ-MOD output in Fig. 1 is an analog signal wave which deals with the pass band signal with a carrier frequency of f_c in the continuous time domain. Since f_c in general is much higher than $1/T_s$, a continuous-time base band OFDM signal with the symbol period T_s and the corresponding pass band signal with the carrier frequency f_c have almost the same waveform. However, in general, the waveform for the discrete-time baseband signal may not be the same as that for the continuous-time baseband signal. In practice, the PAPR for the continuous-time baseband signal can be measured only after implementing the actual hardware, including digital-to-analog convertor (DAC).

In other words, measurement of the PAPR for the continuous-time baseband signal is not straightforward. Therefore, there must be some means of estimating the PAPR from the discrete-time signal. Fortunately, it is known that the discrete-time baseband signal can show almost the same PAPR as the continuous-time baseband signal if it is four times interpolated (oversampled) [9].

The first task to consider is that the CAZAC-OFDM spectrum of 1024 subcarriers with the bandwidth (BW) of 18MHz is centered on the carrier frequency of f_c, 27MHz bandwidth to the left, i.e., the space of 1536 null subcarriers, and in the same way 27 MHz bandwidth to the right. A block diagram of the generation of CAZAC-OFDM symbols is shown inFig.6.

The second step to produce the signal is to apply a transmit filtering with proper roll-off by a FFT-window processing, and to avoid aliasing by low-pass filtering.

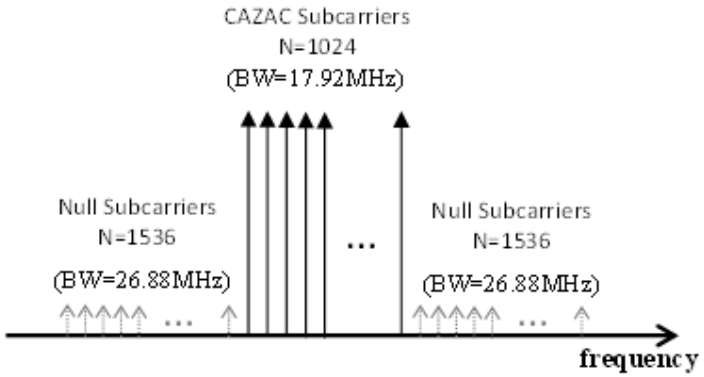


Figure 5. Subcarrier allocation based on four times interpolated (oversampled) FFT processing.

Table I. Major Simulation Parameters

Number of CAZAC Subcarriers	1024 [Data: 1019, Pilot: 5]
IFFT & FFT size	4096
Symbol Period T_s	57 μ sec
Guard Interval	5 μ sec
Modulation of Subcarriers	16QAM
Data Rate	65.74 Mbps
Channel Model	AWGN Flat Rayleigh Fading -Doppler Frequency ($f_D T_s$): 8.6×10^{-4} Frequency Selective Fading -Notch Filter Center frequency 0Hz Bandwidth 70kHz Gain -7dBm

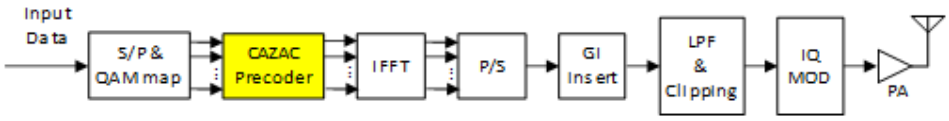


Figure 6. The CAZAC-OFDM transmitter based on four times interpolated (oversampled) FFT processing.

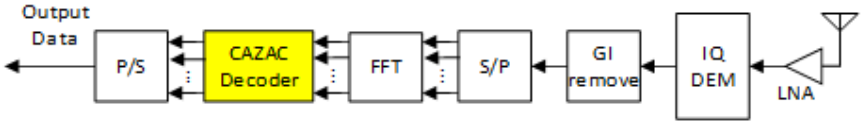


Figure 7. The receiver block diagram for spectral controlled CAZAC-OFDM signal.

3.2. Simulation Results

We carried out our simulation with the MATLAB/Simulink to evaluate the performance of the system. Table I summarizes the simulation parameters.

Figure 8 shows the spectrum for the 16QAM CAZAC-OFDM system when oversampling and filtering are performed. It is confirmed that the spectral performance satisfies the 3GPP LTE specifications.

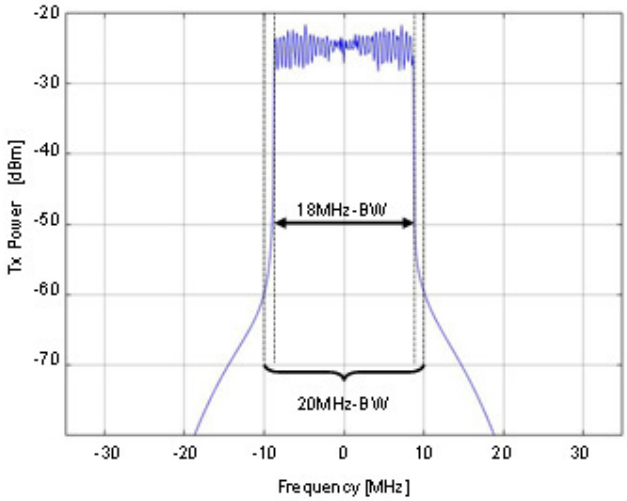


Figure 8. The spectrum controlled 16QAM CAZAC-OFDM signal.

4. PAPR CONTROL

4.1. Procedure of PAPR Control

The CAZAC precoding scheme makes the PAPR of M-array quadrature amplitude modulation (M-QAM) OFDM signals into the PAPR of M-QAM single-carrier signals [7]. Therefore, it dramatically improves the PAPR of OFDM signals. However, an aggressive spectral control results in a high level of correction required in the out of band region, resulting in greater peak regrowth and lead to unacceptable degradation of the PAPR [9].

Figure 9 shows the amplitude of band pass and four times oversampled QPSK CAZAC-OFDM signal with subcarrier allocation shown in Fig. 5. Here, it is known that actual analog signal, i.e., continuous-time signal, can be simulated by discrete-time signal with four times oversampling. The original sample points of QPSK are all the same amplitude. However, the interpolation point generated after spectral shaping causes amplitude fluctuation. This means obvious degradation of the PAPR.

One simplest approach of improving the PAPR is to clip amplitude of interpolating points over to a fixed level, because the amplitude of interpolating points has not important information. Correct demodulation of CAZAC-OFDM signals requires only values of the original sample points, and it is possible to skip out the values of the interpolating points. Therefore, if the clipping boundary is greater than or equal to the maximum value of the original sample point, it can be clipped without affecting the BER performance.

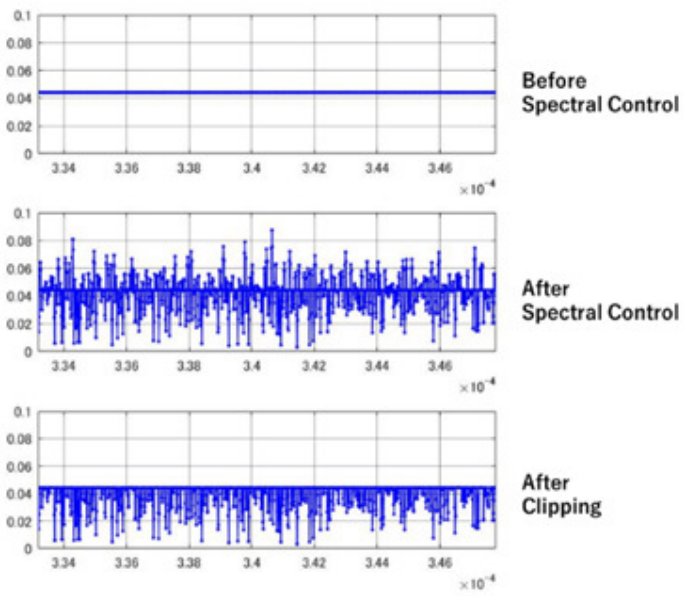
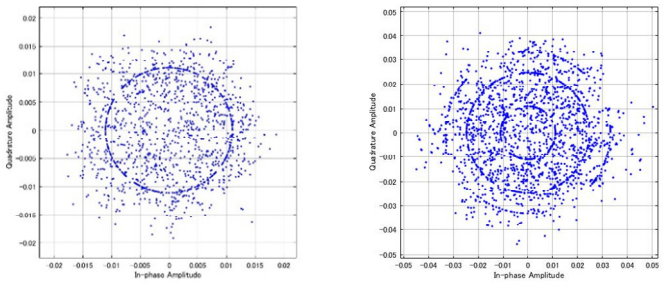
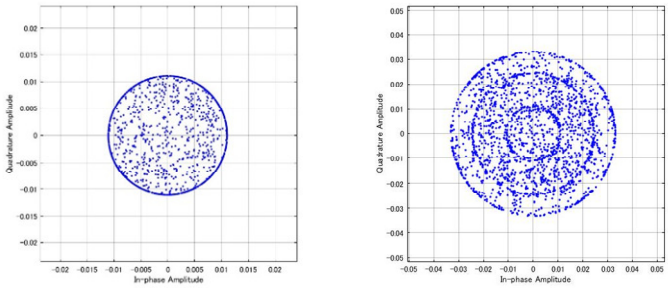


Figure 9. Amplitude of band pass and oversampled QPSK CAZAC-OFDM signal in the time domain.



(a) QPSK (b) 16QAM

Figure 10.The time domain signal points of bandpass and oversampled CAZAC-OFDM signal on the complex phase plane.



(a) QPSK (b) 16QAM

Figure 11.The time domain signal points of band pass and oversampled CAZAC-OFDM signal with clipped amplitude processing.

4.2. Simulation Results

Figure 10 shows the time domain signal points of band pass and four times over sampled CAZAC-OFDM signal with subcarrier allocation shown in Fig. 5. The original sample points of 16QAM CAZAC-OFDM signal are on the three concentric circle orbits as shown in Fig. 3 (b). However, many of the interpolating points overfly to outer side of the largest circle orbit. This means obvious degradation of the PAPR.

As noted previously, correct demodulation of CAZAC-OFDM signals requires only values of the original sample points, which are all in the inner-side of the largest circle orbit of 16QAM CAZAC-OFDM signal, and it is possible to clip amplitude of the interpolating sample points. Figure 11 shows the time domain signal points of band pass and oversampled 16-QAM CAZAC-OFDM signal with clipped amplitude processing.

We first considered the complementary cumulative distribution function (CCDF) of PAPR to evaluate the performance of PAPR, which is the probability that PAPR will be higher than a certain PAPR value PAPR0, i.e., Pr (PAPR > PAPR0). Figure 12 plots the CCDF of PAPR of 16QAM CAZAC-OFDM system with and without clipping, as well as the conventional 16QAM OFDM and single-carrier 16QAM systems. We found that the PAPR of the 16QAM CAZAC-OFDM system with clipping was almost equal to that of the single-carrier 16QAM system. Moreover, the PAPR of the 16QAM CAZAC-OFDM system with clipping was improved by 6.3 dB at the CCDF of 10^{-2} compared with the conventional 16QAM OFDM system.

Figure 13 shows spectrum for the 16QAM CAZAC-OFDM system with clipping. The spectral performance looks acceptable although a little degradation is observed. The result indicates the clipping does not bring serious effect to the spectral performance.

We next examined the bit error rate (BER) of the proposed system. We considered the additive white Gaussian noise (AWGN). Figure 14 shows the BER performances of the 16QAM CAZAC-OFDM systems with clipping which are applied for AWGN channels. The results indicate that the CAZAC-OFDM system with clipping does not degrade BER performances.

Then, flat Rayleigh fading was considered as a fading channel and a notch filter was considered as a frequency selective channel. This BER characteristic is shown in Figure 15. For the flat fading channel, CAZAC-OFDM was improved by about 6 dB at the BER of 10^{-3} compared with the conventional OFDM. Also, in the frequency selective channel, a certain error occurs in the conventional OFDM even in a portion with a high SNR. However, in CAZAC-OFDM, since the information of each subcarrier is dispersed throughout the signal by the CAZAC precoder, it has a frequency diversity effect, and even if a specific frequency is lost, the influence becomes small.

Figure 16 shows PAPR vs. Power Added Efficiency in typical PA in the frequency band of 800MHz-2GHz [10]. For a linear amplification, PAs should be operated at the point back off from the saturation point, i.e., the 1dB-gain-suppression-pint to be exact. Here, the value of the back off should correspond to the PAPR of the input signal for a linear amplification. As a large value of the back off brings degradation of PAE, the PAE in the conventional OFDM scheme becomes only 14%. On the other hand, it is expected that the PAE in the CAZAC-OFDM scheme with clipping is 38%. This is comparable value of single carrier signal scheme, i.e., the SC-FDM scheme.

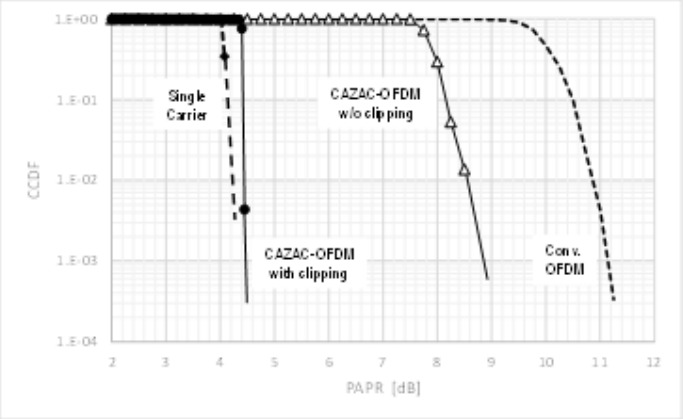


Fig. 12. PAPR of the proposed CAZAC-OFDM.

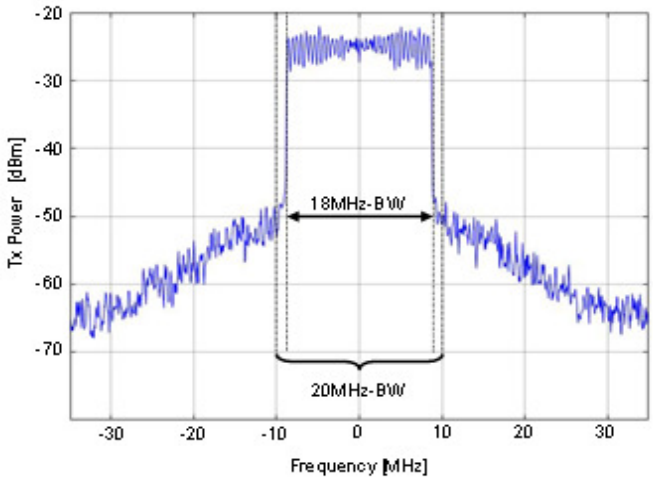


Fig. 13. Spectrum of the 16QAM CAZAC-OFDM Signal with clipping.

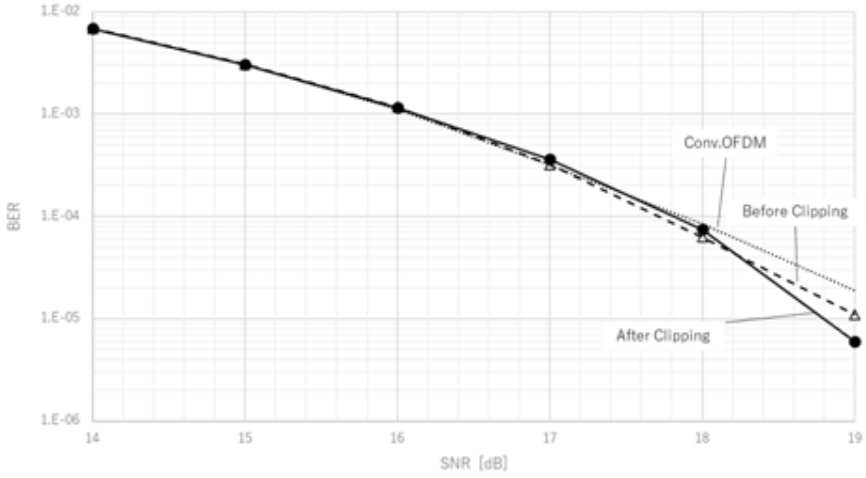


Fig. 14. BER performances of the CAZAC-OFDM under AWGN channel.

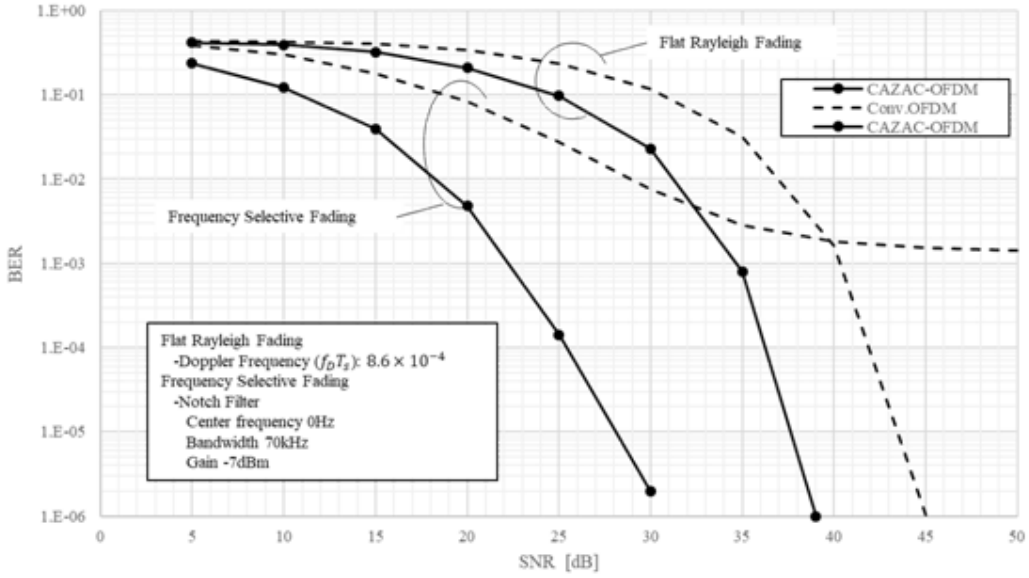


Fig. 15. BER performances of the CAZAC-OFDM under Fading Channel

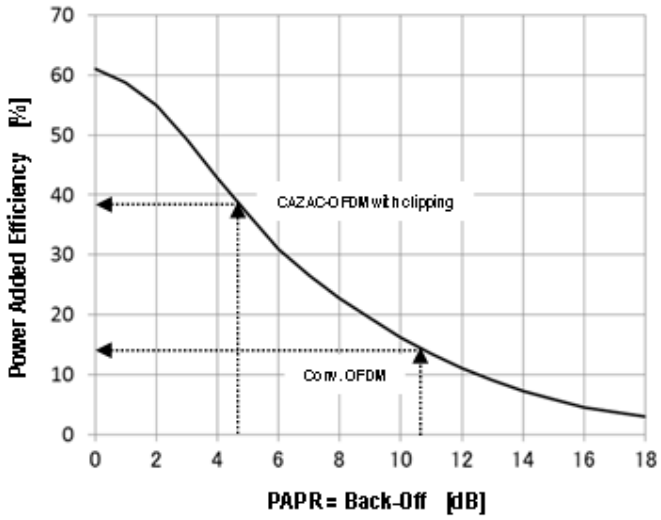


Fig. 16. PAPR vs. Power Added Efficiency of the PA[10].

Table II. Comparisons of PAPR, PAE and Power Consumption of the PA[10] with each waveform.

Waveform (16QAM)	PAPR [dB] @CCDF=1%	PAPR Reduction [dB]	PAE [%]	Power Consumption @PTx:200mW [mW]
CAZAC-OFDM with clipping	4.5	6.3	38	526
CAZAC-OFDM w/o clipping	8.5	2.3	20	1000
Single Carrier	4.2	6.6	40	500
Conv. OFDM	10.8	-	14	1429

5. CONCLUSIONS

The paper provides available control procedure for PAPR and spectrum managements for the CAZAC-OFDM system. One simplest approach of improving the PAPR is to introduce the four-times oversampling technique and to clip the amplitude of the interpolating points to a fixed level. An accurate clipping boundary is easily defined because we can manage the amplitude of the time-domain signal of CAZAC-OFDM like a single-carrier signal. As the correct demodulation of CAZAC-OFDM signals requires only values of the original signal points, we can improve the PAPR without any degradation of the BER and a little degradation of the spectrum by the clipping. The CAZAC-OFDM scheme provides high-efficiency operation of power amplifiers, i.e., the PAE of 38%, the power consumption of the PAs becomes about one-third of those of conventional OFDM scheme, and it is comparable value of single carrier signal scheme, i.e., the SC-FDM scheme. Furthermore, although the PAPR of conventional OFDM scheme strongly deteriorates as increasing the number of subcarriers, the PAPR of the CAZAC-OFDM scheme does not depend on the number of subcarriers.

Therefore, we believe the CAZAC-OFDM scheme can use for uplink channels of future wireless mobile communication systems.

REFERENCES

[1] IEEE, “Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications,” IEEE Std 802.11-2012 (Revision of IEEE Std 802.11-2007), pp. 1-2695 Mar. 2012.

[2] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, and T. Thomas,“ LTE-advanced: Next-generation Wireless Broadband Technology,” IEEE Wireless Communications, vol. 17, no. 3, pp. 10–22, June 2010.

[3] J. Joung, C. K. Ho, K. Adachi, and S. Sun, “A Survey on Power-Amplifier-Centric Techniques for Spectrum- and Energy-Efficient Wireless Communications,” IEEE Communications Surveys & Tutorials, vol. 17, no. 1, pp. 315–333, Jan. 2015.

[4] H. Seung, Hee and L. Jae, Hong, “An Overview of Peak-to-average Power Ratio Reduction Techniques for Multicarrier Transmission,” IEEE Wireless Communications, vol. 12, no. 2, pp. 56–65, Apr. 2005.

[5] H. Myung, J. Lim, and D. Goodman, “Single Carrier FDMA for Uplink Wireless Transmission,” IEEE Vehicular Technology Magazine, vol. 1,no. 3, pp. 30-38, Sep. 2006.

[6] I. Baig and V. Jeoti, "PAPR Reduction in OFDM Systems: Zad off-Chu Matrix Transform Based Pre/Post-Coding Techniques," in Proc.of the 2nd International Conference on Computational Intelligence, Communication Systems and Networks, pp. 373-377, July 2010.

[7] R. Ishioka, T. Kimura and M. Muraguchi, "A Proposal for a New OFDM Wireless System using a CAZAC Precoding Scheme," Proc. AICT 2017, pp. 47-51, June 2017.

[8] Kazuki Miyazawa, Tomotaka Kimura, Masahiro Muraguchi, "Proposal of visible light OFDM system with CAZAC equalization," 23rd Asia-Pacific Conference on Communications (APCC), pp.491-496, Dec. 2017.

[9] Yoshitsugu Sugai, Yushi Shirato, Tomotaka Kimura and Masahiro Muraguchi, "PAPR and Spectral Control Procedure for OFDM Wireless Systems Using CAZAC Equalization," The Fourteenth Advanced International Conference on Telecommunications (AICT) 2018, pp.75-80, July 2018.

[10] H. Okazaki, A. Fukuda, and S. Narahashi, "Band-reconfigurable high-efficiency power amplifier: 900MHz/1900MHz dual-band PA using MEMS switches," NTT DoCoMo Tech. J., vol. 7, no. 1, pp. 11–18, Jun. 2005.

INTENTIONAL BLANK

OBJECT LOCALIZATION AND ACTIVITIES IDENTIFICATION USING ATTRIBUTE DETAILS IN SMART MEETING ROOMS

Dian Andriana[1,2], Ary Setijadi Prihatmanto[2], Egi Muhammad Idris Hidayat[2], and Carmadi Machbub[2]

[1]Research Center for Informatics, Indonesian Institute of Sciences, Bandung, Indonesia
[2]School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung, Indonesia

ABSTRACT

This paper is concerned with the development of interactive systems for smart meeting rooms. Automated recognition of video events is an important research area. We present an LTL (Linear Temporal Logic) model of basic objects and activities recognition in smart meeting rooms using object attribute details. There are still problems of misrecognizing objects in existing visual recognition methods because lack of enough feature attributive information details. This paper investigates morphological approach to increase recognition accuracy using variability in a limited area of moving object using object attribute details. The proposed methods are also compared to popular and recent methods of visual object and event recognition.

KEYWORDS

Events Recognition and Tracking, Morphological Feature Characteristics

1. INTRODUCTION

This paper is concerned with design of automated interactive systems for smart meeting rooms. Design of automated visual recognition of video events is an important research area to support interactive systems in a very early stage of development. Events and behaviours can be modelled in the context of related visual recognized objects. Previous works [1][2][3] are still unable to visually localize speaker of the meeting rooms and need the help of audio features to identify the speaker. Other researchers [4][5][6] explained smart meeting rooms modelling but did not focused on visual events and behaviours which are important features of smart room systems. This paper aims to focus on visual events or behaviours in the smart meeting room systems including visual localization of the speaker. Research done by [1] uses predefined chair and speaker position. However, in the real situation without predefined scenarios, problems still occur in correctness of object identification due to various object position and illumination [7]. In addition to the face features, the speaker or the participants identification can also be done from other personal attributes such as clothes texture, ties, hair, headscarves, etc. Events and behaviours are also identified visually as the speaker entering the room or approaching the podium, talking, leaving the podium, etc. The participants' interactivity through dialogue can also be identified visually.

Natarajan Meghanathan et al. (Eds) : AIAA, WiMNeT, DPPR - 2018
pp. 63–76, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.81606

The object detection using the popular HOG-SVM (Histogram of Oriented Gradients – Support Vector Machines) method still has problems of correctness due to various object position and illumination [7]. Human visual recognition algorithm has been developed and shows success only in controlled environment featuring face, iris, human action and behaviour [8][9]. Face detection algorithm has been developed and shows success in featuring facial landmarks such as corners of the eyes, the tip of the nose, the mouth, the eyebrows, and the face boundaries using regressing trees [10]. The work describes specific face landmarks in detail features or characteristics such as the eyes, eyebrows, nose, and mouth. However, the facial landmarks method is still not able to describe enough information of facial features compared to human perception capabilities of face recognition. There are still problems with human identification using face and body features while they are moving in pose variation [7]. Many statistical and machine learning methods need huge amount or training data taken from human perception or knowledge about objects [10], but it is still incomplete collected training data [11]. Besides statistical methods, syntactical methods include hierarchical, relational, structural, and morphological methods have been developed for face recognition [12][13][14][15]. Structural hierarchical and relational methods have been done for high level abstraction of image objects modelling, but still have problems for low level implementation [12][13]. Many researchers have developed various human activities and behaviour recognition, such as walking, sitting, bending, and some sport activities, and less work of person identification that does the activities [15].

Personal identification is not only using face, but also can use other personal attributes such as hair, hat, headscarves, body posture, or clothing. Personal attributes can be detected using HOG-SVM, for example the textile textures in clothing. Khan [16] combines colour attributes and shapes and learned using HOG-SVM. Works by Nurhaida [17] use SIFT for static textile images, and less movement of the person wearing the textile. Work by Reddy [18] uses LBP and GLCM combined with KNN and SVM to extract features. Kalantidis [19] uses SIFT and LBP for clothing recognition in different appearances.

In visual object detection, object features has also been investigated through SIFT, SURF, BRISK, BRIEF, FREAK, AKAZE, and ORB [20][21][22][23][24] which are based on corner detector as the best features. They work for simple objects with simple background but fails in more complex objects and complex background. The corner-based methods result in basically random match features which are far from correctness in matching object pairs, and need extended areas from the corner points, for example, using curves for describing shapes. The corner points only do not describe shapes and position of object features.

The popular method HOG-SVM has high accuracy in describing face landmarks using thousands of mean-shapes from provided samples classified by regression tree using fit scores. This provides difference or tolerance values from the mean shapes [10] [25], and able to work with simple shapes / curves, but still have problem with sharp curves / non-simple curves. LBP method [8] can be applied in different size of pixels because of histogram size normalization. However, LBP and HOG-SVM cannot be applied in small size of pixels describing small features. They require more or less similar size of pixels and make distribution of grey level without maintaining the pixel positions. GLCM [18] requires too specific and overfitting pattern with lack of generalization or tolerance due to object position and pose changes. HOG SVM in [10] [25] with high accuracy still has ambiguity results which can be eliminated using additional pixels information processing using pixel grey level curve distance evaluation such as Manhattan distance. Problem with illumination are solved using HOG-SVM, but more precise result revalidated using grey level pixel values. Solving ambiguity in HOG-SVM method basically give primal guidance that all objects can be detected using provided samples by their appearances.

Related objects can elaborate more complex problem in scene understanding in wider context of objects, for example in recognition of events, activities or behaviours of moving human objects,

such as human activities or behaviours. In this paper we use lecture and conference activities in a smart meeting room as a study case of events recognition. Images processed from videos which are sampled from courtesy of Youtube [26][27].

2. MODELLING BASIC OBJECTS AND ACTIVITIES OF SMART MEETING ROOM

Personal visual recognition performance using face and body postures can decrease because of movement in different poses [7]. To better recognized and discriminate personal face and body characteristics, we can use attributes of the human object such as the eyes, nose, mouth, hair, headscarf, hat, ties, fabric cloth texture, etc. [28][29][30][31][32][33]. After personal recognition, related events should be recognized in videos by tracking consecutive events or activities. For events or activities recognition, formal models are used for reasoning. In paper by [34] video events are modelled using Petri nets and formalized using LTL (Linear Temporal Logic) formulas to provide guidance for programming implementation.

Basic activities that can be recognized visually in lecture / conference activities in the meeting room include the speaker identification using attributes details, different attributes of the participants, the speaker and the participant activities, layout of display, podium, tables and chairs related to speaker and participants position. This paper focuses on visual recognition without neither predefined position of the camera nor the speaker. Room layouts such as round tables arrangements as recognized objects are included in the model, for example it is useful for controlling dynamic movement of cameras or microphones in smart rooms. The speaker recognition is the central of the system, and recognized not only from the face, but also from detail attributes and behaviours. In this paper we focus on detail attributes, behaviours of the speaker and the participants of the conference, and the room layouts. For example, we use semi round tables layout as depicted in Figure 1, and we model the semi round tables layout in LTL formulae as follows.

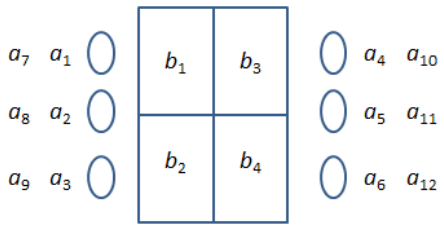


Figure 1. An example of semi round tables arrangement

$$\mathbf{GF}s \tag{1}$$

$$s \vDash \mathbf{F}a_1 \wedge \mathbf{F}a_2 \wedge \mathbf{F}a_3 \ldots \wedge \mathbf{F}a_{12} \tag{2}$$

where

s = satisfied condition of semi round tables arrangement
a1 = the first position of detected person / chair facing right
a2 = the second position detected person / chair facing right
a3 = the third position detected person / chair facing right
a4 = the fourth position of detected person / chair facing left
a5 = the fifth position detected person / chair facing left

a6 = the sixth position detected person / chair facing left
a7 = the first position of detected person / chair facing table
a8 = the second position of detected person / chair facing table
a9 = the third position of detected person / chair facing table
a10 = the fourth position of detected person / chair facing table
a11 = the fifth position detected person / chair facing table
a12 = the sixth position detected person / chair facing table
b1= b2 = b3 = b4 = four small tables
or replaced by b5= one big table.

$$a_4 \vDash (\mathbf{F}a_{41} \wedge \mathbf{F}a_{42} \wedge \mathbf{F}a_{43}) \vee (\mathbf{F}a_{44} \wedge \mathbf{F}a_{45} \wedge \mathbf{F}a_{46}) \quad (3)$$

where

a_{41} = x-coordinate left position of the person / chair a_1< x-coordinate left position of the table b_1
a_{42} = y-coordinate top position of the person / chair a_1≥ y-coordinate top position of the table b_1
a_{43} = y-coordinate bottom position of the person / chair a_1≤ y-coordinate bottom position of the table b_2
a_{44} = x-coordinate left position of the person / chair a_1< x-coordinate left position of the table b_5
a_{45} = y-coordinate top position of the person / chair a_1≥ y-coordinate top position of the table b_5
a_{46} = y-coordinate bottom position of the person / chair a_1≤ y-coordinate bottom position of the table b_5

We define an example of a speaker through talking activity and tie attribute as follows.

$$\Omega = \text{count(talk)} \geq \epsilon \quad (4)$$

$$\neg\, \Omega = \text{count(talk)} < \epsilon \quad (5)$$

$$\Psi = (\mathbf{G}t \wedge (\Omega \mathbf{R}\, d)) \quad (6)$$

where t = the person is wearing a tie and d = counting mouth movement (talking) as described in Figure 2. $\Omega$is a condition when the counting d exceeds a number ϵ which is an assumption of a number of mouth detected opening which is also based on position differences between detected upper and lower lips from theprevious frame-to-frame movement, and also based on counting of the second order differences of the position differences between the detected upper and lower lips as a representation of talking mouth movement. In equation (7) and Figure 3, we also define an example of an interactive participant which does not have a speaker attributes, probably raises hand and stands up for question talks.

$$\mathbf{G}\neg\psi \wedge \text{Fh} \wedge \mathbf{X}((\mathbf{F}c) \vee (\Omega \mathbf{R}\, d)) \quad (7)$$

where

$\neg\psi$ = that person is not a speaker (maybe a participant)
c= the person is standing up (probably eventually)
d = counting mouth movement (talking)
Ω = count(talk) ≥ϵ_1
h= the person raise hand

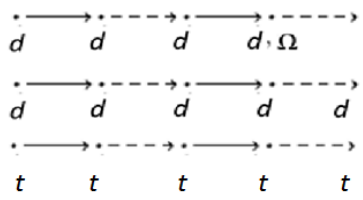


Figure 2. LTL diagram of talking counting and tie attribute

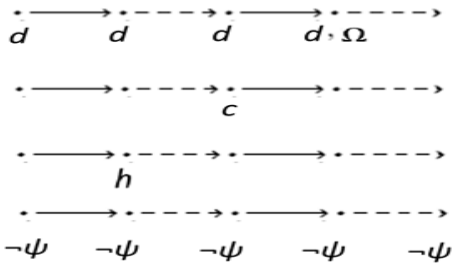


Figure 3. LTL diagram of participant interactivity

In Figure 4, we define an example of a speaker detected with tie attribute and talking activities. We also define other general personal detail attributes such as hair, hat, headscarf, and clothing textile patterns.

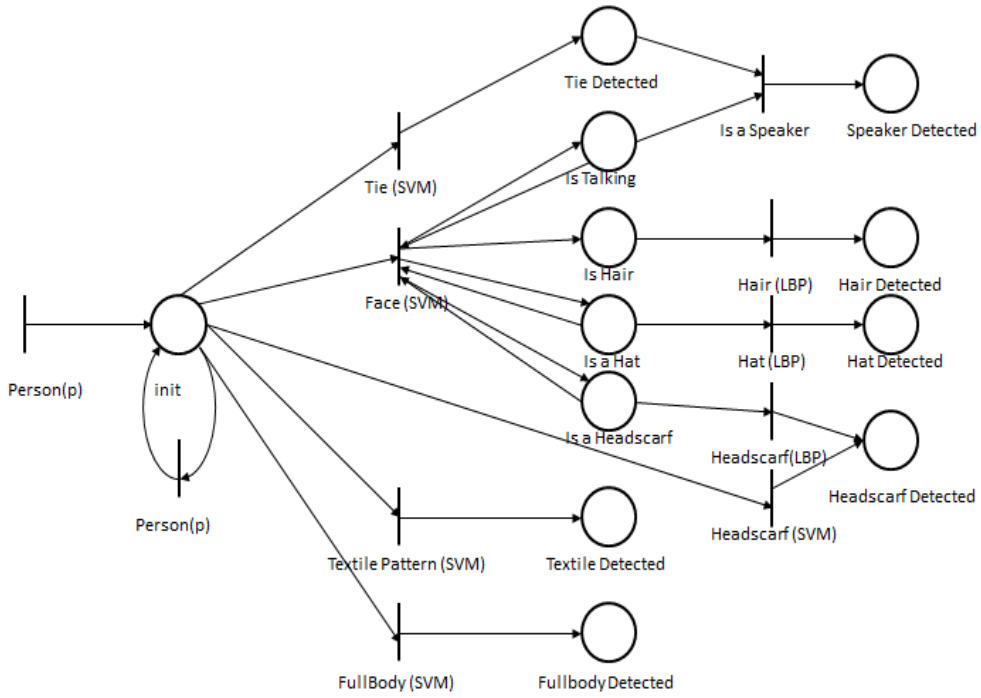


Figure 4. Petri nets diagram of personal detail detection

In Figure 5, we define a scenario of a person movement towards a detected podium, being near podium, talks, and then move away after finish talking. Figure 6 and Figure 7 describe in further details about a person movement detected from position coordinate changes and being near podium coordinates, and eventually detected as a speaker.

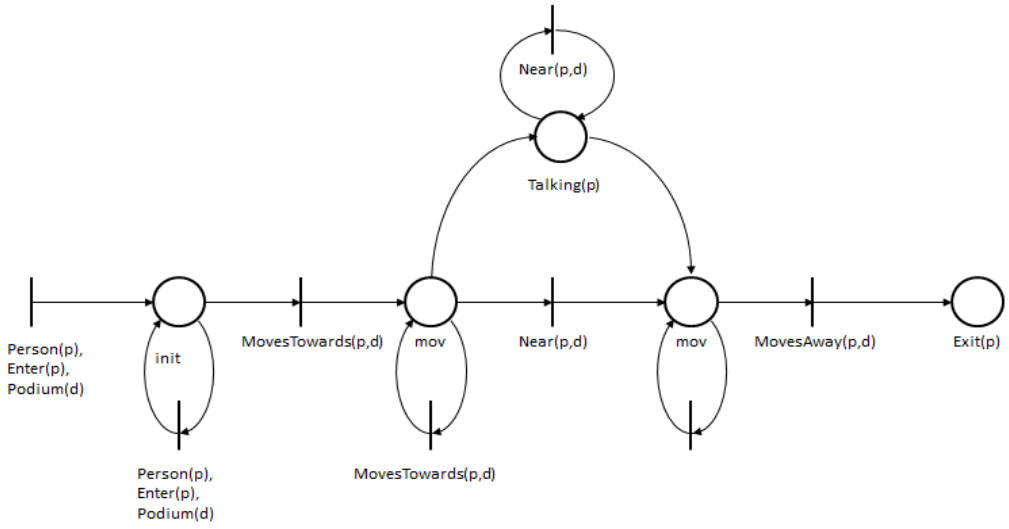


Figure 5. Petri nets diagram of speaker scenario

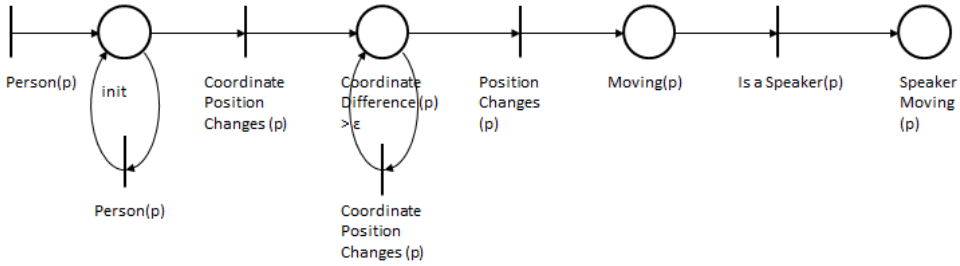


Figure 6. Petri nets diagram of moving speaker

Figure 8 also describes a speaker detected from collected score of checking detected tie, being near podium coordinates, and talking activities from detected lips moving coordinates.

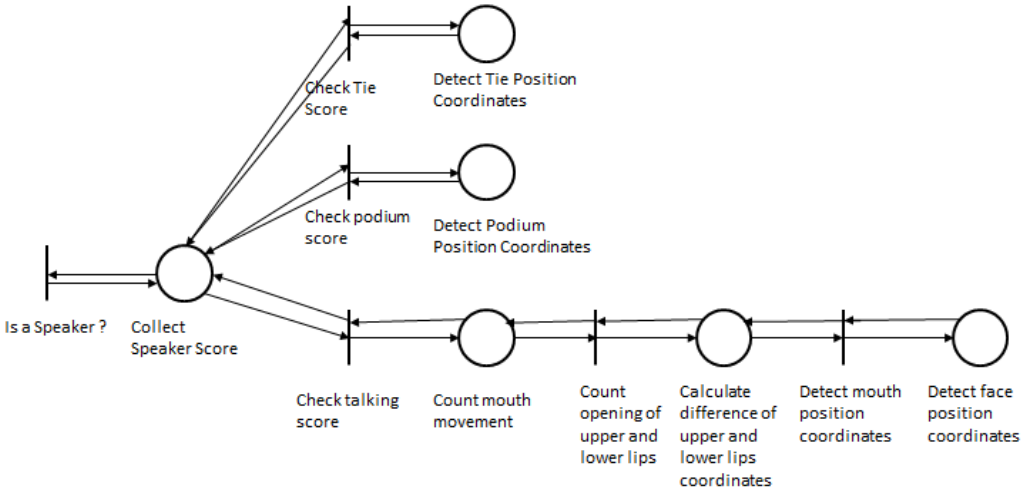


Figure 7. Petri nets diagram of speaker near podium

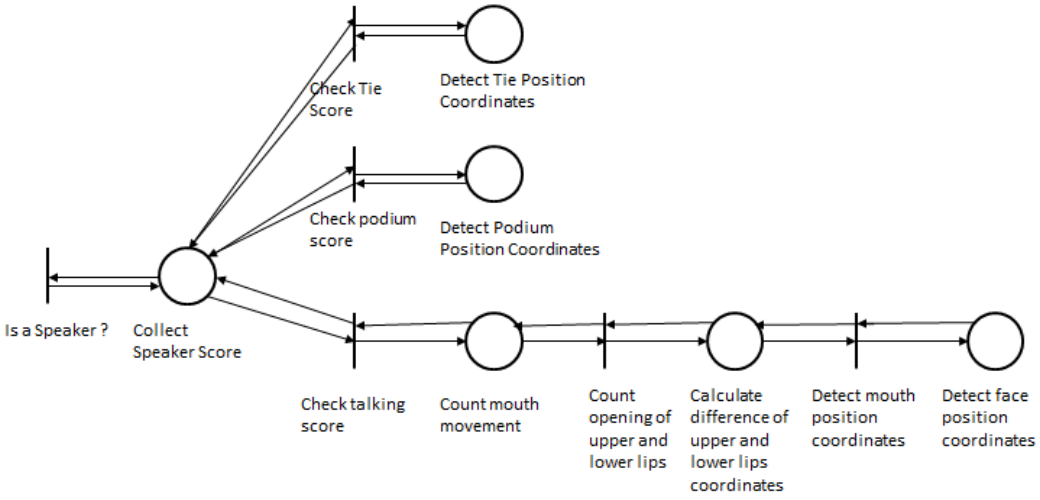


Figure 8. Petri nets diagram of defining speaker properties

3. DETAIL ATTRIBUTE DETECTION USING GRADIENTS

We use an alternative of human recognition to recognize hair texture on moving person in the smart meeting room. In case of small size of the object region of interest (ROI), for example for far objects from camera which implies smaller size of ROIs, we cannot use HOG-SVM due to size of small size area. We cannot also use LBP due to value similarities which are not able to distinguish true and false features. We use Multiple Linear Regression formulas in [31] to describe feature of objects. Examples for identification of the speaker's white hair texture distinguished from the audience's brown hair with guidance of face landmark position are shown in Figure 9 and Figure 10. With position guidance of face landmarks and HOG-SVM object learning [25][35] detection area, we pick hair texture samples from one frame in video and compare them to the other frames. We formulate the pixel grey level value differences below, and then calculate a mean value between Q and R.

$$C_3 < |Q| < C_4\,,\ Q = \{df_t | df_t = f_t - f_{t+1} \cap df_t < 0\} \tag{8}$$

$$C_5 < |R| < C_6\,,\ R = \{df_t | df_t = f_t - f_{t+1} \cap df_t > 0\} \tag{9}$$

The step in Q and R sets are getting the optimum number of pixel points in certain ranges which have negative (or positive) gradients $df_t$with their successor points in the curve of pixel gray levels.

In Table I we use the formulas to describe similarity between sample and tested object features in moving person. We use three ROIs to enable variability of feature areas in moving object, but still in guidance of face landmarks. For example, experiment 1 ROI 2 shows a minimum mean value of pixel gradients Q and R values, and all other experiments in Table I also shows a minimum value at least on an ROI, which is an expected result of the mean formula. In the other side, Manhattan distance calculation fails to show consistent behavior while giving lower value in false features in Table II experiment 4 ROI 1 and higher value in true features in Table I experiment 5 ROI 2. In correct location area, the distance values are expected to be consistently lower in true features and higher in false features. The true and false features are then validated by human eyes, and the mean formula is expected to show consistent expected results.

Table 1. Distance measurement comparison examples for true features.

Experiment	ROI	Mean	Manhattan
1	1	38	2454
	2	5	1301
	3	7	4082
2	1	38	2454
	2	7	1340
	3	3	4092
3	1	13	2688
	2	11	4506
	3	6	4520
4	1	5	2932
	2	10	4412
	3	2	4614
5	1	81	41662
	2	6	36304
	3	41	17976
6	1	73	41640
	2	17	33918
	3	72	21699

Table 2. Distance measurement comparison examples for false features.

Experiment	ROI	Mean	Manhattan
1	1	105	45000
	2	100	46247
	3	85	45045
2	1	105	45027
	2	98	46287
	3	92	45432
3	1	106	45423
	2	104	46321
	3	101	44936
4	1	78	15647
	2	23	29486
	3	40	33221
5	1	82	16664
	2	22	29313
	3	33	32859
6	1	242	66257
	2	118	121455
	3	138	158312

Figure 9. Speaker white hair texture feature, picture is courtesy of Youtube [26]

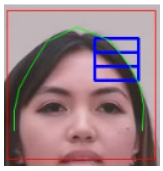


Figure 10. Audience hair texture feature, picture is courtesy of Youtube [26]

4. RESULTS AND DISCUSSIONS

In Table 1 and Table 2, we use mean measurement formula (8) and (9) to clearly distinct true features from false features. Table 1 show mean values on each ROI for each experiment, and for each experiment there exist minimum mean values which are less than 20 for true features and more than 20 for false feature. The minimum mean values occur on at least one of the ROIs of each experiment of true features. In the other side, Manhattan distance measurement cannot distinguish clearly between true and false features.

Choices of segmentation methods and ROI areas are important for defining detail features of object detection. ROI area variability can be extended with guidance of other relatively more reliable features such as face landmarks. There are also possibilities of combining criteria with other attributes such as a person wearing ties. Another challenge is noisy background of the objects. The detection methods can be combined with other features such as colour, but colour attribute is not always useful in outdoor sunny daylight environment which may affect brighter object colour.

Basic objects and activities in smart meeting room are also as the results of objects and activities detection using the LTL formulas and petri nets implementation. Figure 11 shows speaker detection through tie attributes, mouth movement counting, and white hair property in ROIs. In this figure, the mean formula shows white hair property similarity as the lowest value 3 in the bottom ROI position, when Manhattan formula shows inconsistently higher value 4449 compared to 2630 and 1381, also when LBP formula shows indistinctive all zero values in all the ROIs. Figure 12 shows an audience with no tie attributes, and no white hair property since the lowest mean value is 10 which is higher than 3 as the value of white hair property, when Manhattan formula also shows inconsistently higher value compared to other ROIs, and LBP formula shows errors. Figure 13 shows moving person detected, and speaker detected from clothing pattern in the middle ROIs which has the lowest mean value, but LBP formula shows inconsistent higher value in the middle ROI. Finally, in Figure 14 LTL formulas can determine a speaker movement and positioned near podium which is also previously detected by HOG-SVM object detection. Figure 14 also shows the lowest mean value in the bottom ROI, when LBP shows inconsistent higher value in the bottom ROI. Figure 15 is demonstrated by Yolo [36] object detection with correct position of a person, chairs, and a table, but it still misrecognizes a working desk as a dining table. Figure 15 also shows indistinctive similar LBP values for all three ROIs, when the mean value shows the lowest value in the bottom ROI for clothing pattern similarity. The mean formula can distinct object properties in some cases better than HOG and LBP because it maintains the pixel positions rather than distribution of pixel values, also the mean formula works better than HOG and LBP on small size of picture patches as the object properties.

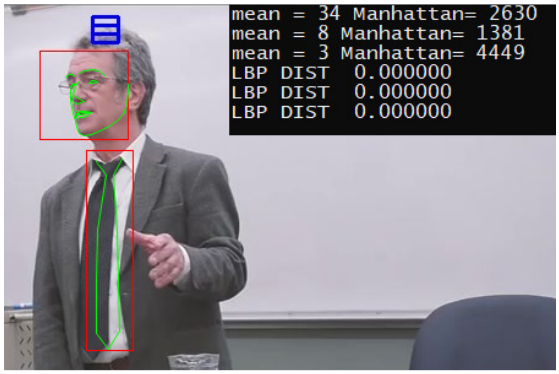


Figure 11. Speaker detected from tie attributes, mouth movement counting, and white hair property in one of the three ROIs, picture is courtesy of Youtube [26]

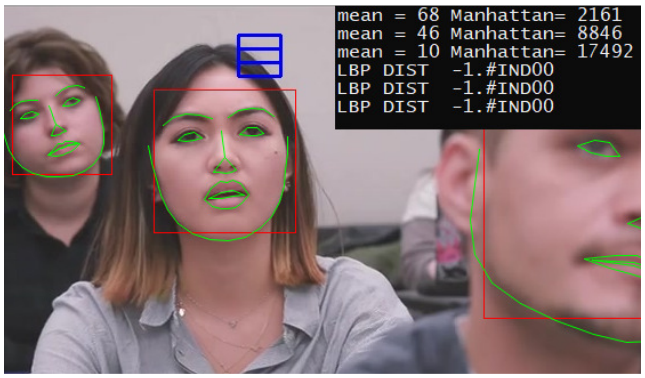


Figure 12. Participant detected from no tie attribute, mouth movement counting, and no white hair property in one of the three ROIs, picture is courtesy of Youtube [26]

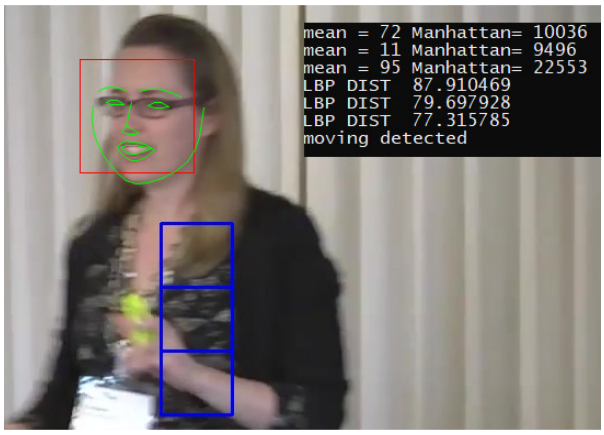


Figure 13. Moving person detected, and speaker detected from clothing pattern in one of the three ROIs, picture is courtesy of Youtube [26]

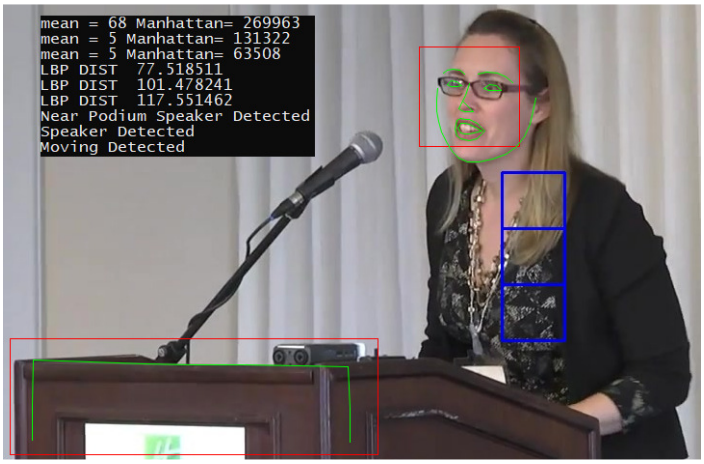


Figure 14. Moving speaker detected near podium, and speaker detected from clothing pattern in one of the three ROIs, picture is courtesy of Youtube [27]

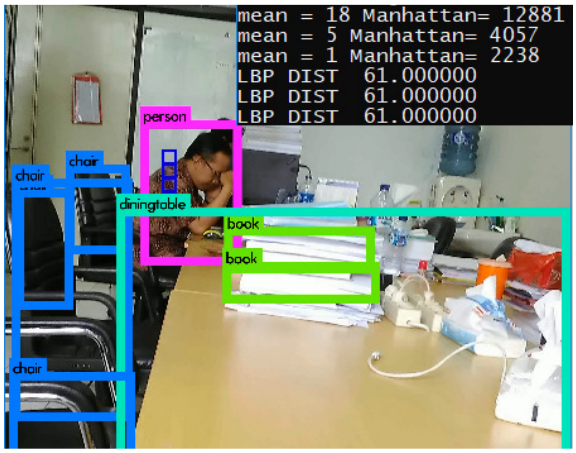


Figure 15. Person, table and chair detected positions using Yolo [36], and a person detected from clothing pattern in one of the three ROIs

5. CONCLUSIONS

This paper has shown a LTL model and its implementation on visual objects and activities recognition in smart meeting rooms. We have also shown that optimum number of pixel points in certain ranges which have negative (or positive) gradients of grey level pixel values can be used to distinct features to increase attribute detail recognition. There are still challenges to distinct more features to increase accuracy of moving object and event recognition. Another possibility is using variability of the gradient sequence.

ACKNOWLEDGEMENTS

We would like to thank Institute of Technology Bandung, Indonesian Institute of Sciences, Indonesian Ministry of Research, Technology and Higher Education especially Saintek Scholarship program to support this research.

REFERENCES

[1] Andrey Ronzhin, Alexander Ronzhin,& Viktor Budkov, (2011) "Audiovisual Speaker Localization in Medium Smart Meeting Room", 2011 8th International Conference on Information, Communications & Signal Processing, Singapore, pp1-5.

[2] Zhao Li, Thorsten Herfet, Martin Grochulla,& Thorsten Thormählen, (2012) "Multiple Active Speaker Localization Based on Audio-Visual Fusion in Two Stages", IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), Hamburg, pp1-7.

[3] Ivan Galov, Rustam Kadirov, Andrew Vasilev & Dmitry Korzun, (2013) "Event recording in Smart Room", 13th Conference of Open Innovations Association (FRUCT), Petrozavodsk, pp20-28.

[4] L. D. Tran et al, (2016) "A Smart Meeting Room Scheduling and Management System with Utilization Control And Ad-Hoc Support Based On Real-Time Occupancy Detection", IEEE Sixth International Conference on Communications and Electronics (ICCE), Ha Long, pp186-191.

[5] Simon Mayer, Nadine Inhelder, Ruben Verborgh, Rik Van de Walle, &Friedemann Mattern, (2014) "Configuration of smart environments made simple: Combining visual modeling with semantic metadata and reasoning"; International Conference on the Internet of Things (IOT), Cambridge, MA, pp61-66.

[6] Maik Wurdel, (2011) "An Integrated Formal Task Specification Method for Smart Environments"; Logos Verlag.

[7] Vishal M. Patel, Jaishanker K. Pillai, & Rama Chellappa, (2011) "Image and Video-Based Biometrics"; T.B. Moeslund et al. (eds.), Visual Analysis of Humans, Springer-Verlag London, pp437-454.

[8] Timo Ahonen, Abdenour Hadid, &Matti Pietikainen, (2004) "Face Recognition with Local Binary Patterns," Computer Vision – ECCV, pp469–481.

[9] Javier Galbally, Sebastien Marcel, &Julian Fierrez, (2014) "Biometric Antispoofing Methods: A Survey in Face Recognition", IEEE Access,Vol. 2, pp1530-1552.

[10] Vahid Kazemi& Josephine Sullivan, (2014) "One Millisecond Face Alignment with an Ensemble of Regression Trees", CVPR '14 Proc. of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp1867-1874.

[11] Stefanos Zafeiriou, Cha Zhang, &Zhengyou Zhang, (2015) "A Survey on Face Detection in The Wild: Past, Present And Future",Comput. Vis. Image Underst., vol. 138, pp1–24.

[12] Laura Antanas, Paolo Frasconi, Fabrizio Costa, Tinne Tuytelaars, & Luc De Raedt, (2012) "A Relational Kernel-based Framework for Hierarchical Image Understanding", SSPR'12 / SPR'12 Proc. of the 2012 Joint IAPR Int. Conf. on Structural, Syntactic, and Statistical Pattern Recognition, pp171-180.

[13] Laura Antanas, Martijn van Otterlo, José Oramas M., Tinne Tuytelaars & Luc De Raedt, (2012) "A Relational Distance-based Framework for Hierarchical Image Understanding", Int. Conf. on Pattern Recognition Applications and Methods, Vol. 2, pp206–218.

[14] Robert Olszewski, (2001) "Generalized Feature Extraction for Structural Pattern Recognition in Time-Series Data", Pittsburgh PA: Carnegie Mellon University.

[15] Hua-Tsung Chen, Yu-Zhen He, Chun-Chieh Hsu, Chien-Li Chou, Suh-Yin Lee, &Bao. Shuh P. Lin, (2014) "Yoga Posture Recognition for Self-training", Int. Conf. on Multimedia Modeling (MMM 2014), pp496-505.

[16] Fahad Shahbaz Khan, Rao Muhammad Anwer, Joost van de Weijer, Andrew D. Bagdanov, Maria Vanrell, & Antonio M. Lopez, (2012) "Color Attributes for Object Detection", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp3306-3313.

[17] Ida Nurhaida, Ary Noviyanto, Ruli Manurung, &Aniati M. Arymurthy, (2015) "Automatic Indonesian's Batik Pattern Recognition Using SIFT Approach",Procedia Computer Science, Vol. 59, pp567-576.

[18] R. Obula Konda Reddy, B. Eswara Reddy, & E. Keshava Reddy, (2014) "An Effective Glcm And Binary Pattern Schemes Based Classification for Rotation Invariant Fabric Textures",International Journal of Computer Engineering Science (IJCES), Vol. 4 Issue 1.

[19] Yannis Kalantidis, Lyndon Kennedy, & Li-Jia Li, (2013) "Getting the Look: Clothing Recognition and Segmentation for Automatic Product Suggestions in Everyday Photos", In Proceedings of the 3rd ACM conference on International conference on multimedia retrieval (ICMR '13). ACM, New York, NY, USA, pp105-112.

[20] Herbert Bay, Tinne Tuytelaars, & Luc Van Gool, (2006) "Surf: Speeded Up Robust Features", Computer Vision–ECCV, Springer Berlin Heidelberg, pp404-417.

[21] Stefan Leutenegger, Margarita Chli, & Roland Y. Siegwart, (2011) "BRISK: Binary Robust Invariant Scalable Keypoints", Computer Vision (ICCV), IEEE International Conference on. IEEE.

[22] Alexandre Alahi, Raphael Ortiz, &Pierre Vandergheynst, (2012) "Freak: Fast Retina Keypoint", Computer Vision and Pattern Recognition (CVPR), IEEE Conference on. IEEE.

[23] Michael Calonder, Vincent Lepetit, Cristoph Strecha, & Pascal Fua, (2010) "Brief: Binary Robust Independent Elementary Features", Computer Vision–ECCV, Springer Berlin Heidelberg, pp778-792.

[24] Ethan Rublee, Vincent Rabaud, Kurt Konolige, & Gary Bradski, (2011) "ORB: an efficient alternative to SIFT or SURF", Computer Vision (ICCV), IEEE International Conference on. IEEE.

[25] Pedro F. Felzenszwalb, Ross B. Girshick, David McAllester, & Deva Ramanan, (2010) "Object Detection with Discriminatively Trained Part Based Models",IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 32, No. 9.

[26] https://www.youtube.com/watch?v=bGP9JRhQXak

[27] https://www.youtube.com/watch?v=Yo845WG3KH8&t=173s

[28] Dian Andriana, (2013) "Linear Function and Inverse Function with Weight Ratio for Improving Learning Speed Of Multi-Layer Perceptrons Feed-Forward Neural Networks", Proc. IC3INA, pp255-259.

[29] Dian Andriana, (2015) "Contiguous Uniform Deviation for Artificial Neural Network Pattern Recognition",Advanced Science Lett. Vol. 21 No. 2, pp189-191.

[30] Dian Andriana, Carmadi Machbub, &Ary Setijadi Prihatmanto Prihatmanto, (2015) "Opponent Zigzag Movement Model Capture and Prediction in Robotic Soccer", International Conference on Interactive Digital Media.

[31] Dian Andriana, Ary Setijadi Prihatmanto, Egi Muhammad Idris Hidayat, &Carmadi Machbub, (2017) "Contiguous Uniform Deviation for Multiple Linear Regression in Pattern Recognition",Journal of Physics: Conference Series, 801 (1), 012046, 2017.

[32] Dian Andriana, Ary Setijadi Prihatmanto, Egi Muhammad Idris Hidayat, &Carmadi Machbub, (2017) "Combination of Face and Posture Features for Tracking of Moving Human Visual Characteristics",International Journal on Electrical Engineering and Informatics, Vol. 9, Iss. 3, pp616-631.

[33] Gabriela E. Martinez, Olivia Mendoza, Juan R. Castro, A. Rodriguez-Diaz, Patricia Melin, & Oscar Castillo, (2015) "Response Integration in Modular Neural Networks Using Choquet Integral with Interval Type 2 Sugeno Measures", Fuzzy Inf. Processing Society (NAFIPS) held jointly with 2015 5th World Conf. on Soft Comput. (WConSC), 2015 Annual Conf. of the North American, Redmond WA, USA: IEEE, pp1-6.

[34] Piotr Szwed & Mateusz Komorkiewicz, (2013) "Object Tracking and Video Event Recognition withFuzzy Semantic Petri Nets", 2013 Federated Conference on Computer Science and Information Systems, Kraków, pp167-174.

[35] Davis E. King, (2009) "Dlib-ml: A Machine Learning Toolkit",Journal of Machine Learning Research 10, pp1755-1758.

[36] Joseph Redmon, Santosh Divvala, Ross Girshick & Ali Farhadi, (2016) "You Only Look Once: Unified, Real-Time Object Detection", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, pp779-788.

AUTHORS

DianAndriana is now a PhD student in School of Electrical Engineering and Informatics, the Institut Teknologi Bandung (ITB), where she also completed her bachelor' and master's degree. She is also a researcher at the Research Center of Informatics of the Indonesian Institute of Sciences. Her researches interests include decision support and intelligent systems. She has 5 papers published in Scopus indexed journals and conferences.

Ary Setijadi Prihatmanto graduated with B.E. and M.S. in Electrical Engineering at Institut Teknologi Bandung in 1995 and 1998, and received his PhD in Applied Informatics from Johannes Kepler University of Linz, Austria in 2006. He is an associate professor & lecturer of School of Electrical Engineering & Informatics, Institut Teknologi Bandung since 1997. He is also the president of Indonesia Digital Media Forum since 2009. His main interests are Human-Content Interaction, Computer Graphics & Mixed-Reality Application, Machine Learning & Intelligent System, Intelligent Robotics, and Cyber-Physical System.

Egi Hidayat received the B.Eng. degree in Electrical Engineering from the Institut Teknologi Bandung (ITB) in 2003, M.Sc. degree in Control and Information Systems from University of Duisburg-Essen in 2007, and Ph.D. degree in Electrical Engineering from Uppsala University in 2014. He is now with the Control and Computer Systems Research Division at School of Electrical Engineering and Informatics, ITB. His research interests are mainly within the area of system identification and signal processing.

Carmadi Machbub got Bachelor degree in Electrical Engineering from the Institut Teknologi Bandung (ITB) in 1980, Master degree (DEA) in Control Engineering and Industrial Informatics in 1988, and Doctorat degree in Engineering Sciences majoring in Control Engineering and Industrial Informatics from Ecole Centrale de Nantes in 1991. He is now Professor and Head of Control and Computer Systems Research Division, School of Electrical Engineering and Informatics, ITB. His current research interests are in machine perception and control.

ENSEMBLE LEARNING BASED VOTING MODEL FOR DYNAMIC PROFILE CLASSIFICATION AND PROJECT ALLOTMENT

Suhas Tangadle Gopalakrishna and Vijayaraghavan Varadharajan

Infosys Limited, Bengaluru, India

ABSTRACT

Every year, lakhs of students right from college enter professional life through various recruitment activities conducted by the organization. The allotment of projects to the new recruits, carried out by the HR team of the organization is usually a manual affair. It is a time consuming and a tedious process as it involves manually opening each resume and analysing it one by one in order to assign a project. Companies round the globe are leveraging the power of artificial intelligence and machine learning to increase their productivity. In this paper, we present one such use case wherein artificial intelligence is leveraged by the organisation in allotment of projects to the new recruits. Current machine learning tools help in the allotment of projects to a few known popular domains on which the classifier has been trained explicitly. We tackle the problem with an ensemble learning based voting classifier consisting of 5 individual machine learning classifiers, voting to classify the profile of the candidate into the relevant domain. The knowledge extracted from the profiles for which there is no majority consensus among the individual classifiers is used to retrain the model. The proposed model achieves a higher accuracy in classifying resumes to proper domains than a standard machine learning classifier which is solely dependent on the training set for classification. Overall, emphasis is laid out on building a dynamic machine learning automation tool which is not solely dependent on the training data in allotment of projects to the new recruits.

KEYWORDS

Ensemble learning based voting classifier, Dynamic classification, Artificial Intelligence, Resume classifier, Association Rule Learning, Latent Dirichlet Allocation

1. INTRODUCTION

The jobs in the IT sector has been growing exponentially ever since it's inception. Companies round the globe recruit thousands of young talent every year. The new recruits have to be allotted projects by the organisation. Usually, HR team is entrusted with the responsibility of allotting the projects to fresh recruits, which is usually a manual affair. Manual allotment of projects to the new recruits by analysing the resumes of the candidates one by one is a tedious and a redundant process. In this paper, we have investigated a mechanism which helps the HR team in allotting the projects to the fresh recruits by considering the skill sets, interests and work experience mentioned in the resume of the candidates. The world of AI has grown significantly in the last decade or so. With the availability of large amounts of data, major tech companies round the globe are leveraging the concept of machine learning to increase productivity. According to recent survey by computer giant IBM [9], there is roughly about 2.5 Exabyte which corresponds

Natarajan Meghanathan et al. (Eds) : AIAA, WiMNeT, DPPR - 2018
pp. 77–88, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.81607

to about 2.5 billion gigabytes (GB) of data in the world. With the advent of big data tools like Hadoop, Spark, companies are capable of storing and analysing large amounts of data to build data hungry machine learning models for the purpose of automation. Enterprises these days are on a spree of expansion in new fields. Hence, they recruit individuals with varied specialization in order to gain traction in different fields. The current machine learning tools for resume classification do not entertain the resumes of candidates with specialization other than a few known, popular fields of domain on which it has been trained on. Thus, they fail to map the profiles with varied specialisation to projects. The main objective of our tool is to solve this issue by introducing the element of dynamism during classification of an instance of resume. This is achieved by building a model which is in a constant learning mode. For those resumes with specializations that has not been included in the training set, the tool extracts the features out of the resume, analyses and passes the knowledge extracted from the profile to a REST API trained on the stack-overflow dump for additional information on the category of domain to which the profile belongs to. Once the related topics based on the results from the Association Rule Mining of the Questions and Answers of the data dump is retrieved, it is analysed and subjected to topic modelling to extract relevant domain for the profile. The specialization of the resume is then added to the training set along with the extracted features. The process of training is then repeated in regular intervals in order to train the classifier on the new specializations. In this way, we eliminate the dependency of the model solely on the training set for categorization of resume and hence, eliminate the limitations of current machine learning models in the market specialized for this purpose.

2. RELATED WORK

Considerable work has been accomplished in the field of text categorization by leveraging the concepts of machine learning in recent times [10]. In [1], the authors compare the different machine learning techniques which can be employed for text categorization [5]. But in this case, the categories into which the text would be classified is specified during the training phase. The model would then classify the text to a domain based on the training data alone. In [2] the authors proposea K-Nearest Neighbours machine learning classifier to classify a text into different clusters. V Ram and Prasanna have highlighted the importance of neural networks in the analysis of textual data [3]. They have argued that by feeding the neural networks with enough examples, the network would be able to predict the category of the text data in the test set more accurately than the conventional machine learning algorithms. Yieng Huang and Jingdeng Chen have argued for a deep learning model to classify text data [4]. Deep learning models which are a special variant of the neural networks discussed above are data hungry. These models have more than one hidden layers compared to conventional neural network models. Previously, research in the field of text categorisation has been solely dependent upon the training data. Our model aims at eliminating the sole dependency on the training set for mapping of profiles to the projects.

3. OUR APPROACH

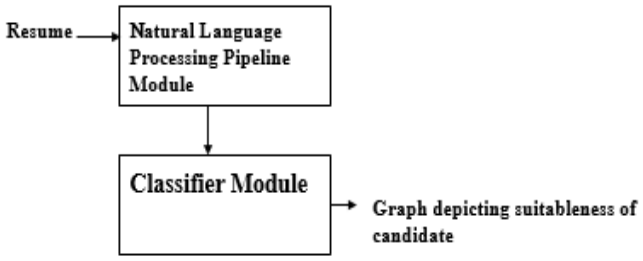


Figure 1. Block Diagram

The application developed comprises of two main modules as shown in Figure 1. The modules are: a) Natural Language Processing Pipeline (NLPP) and b) Classification module. The resumes of the new recruits are fed as an input to the NLPP by the HR team of the organisation. The NLPP module converts the resume into tokens that can be used by the classification module for the actual classification of the resume into a proper domain. The NLPP involves several steps which convert the resume input into tokens. The tokens generated are then fed as an input to the classification module. The classification module analyses the list of tokens generated by the NLPP module in order to allot a domain to the candidate. The application plots a graph depicting the relevance of the candidate with respect to various domains depending upon the knowledge extracted from the resume of the candidate. Hence, instead of only giving details about the most suitable domain for the resume, the relevance of the profile across various domains is plotted in the form of graph. The results from the graph aids the HR team of the organisation in allotting the projects to the candidates, eliminating the tedious process of opening each resume one by one and analysing them manually.

4. NATURAL LANGUAGE PROCESSING PIPELINE

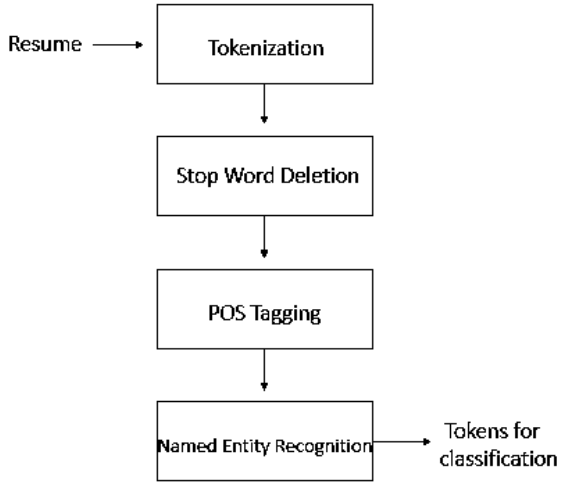


Figure 2: Flow chart of NLPP module

Figure 2 gives a detailed overview of the entire process of NLPP. The new recruits are allotted projects based on the content present in their respective resumes. The data in the resumes of the candidates are subjected to a NLPP in order to obtain only necessary and relevant details. Figure 3 depicts a portion of sample machine learning profile.

- Sentiment analysis and classification using tensorflow in MIT
- Fraud detection using neural networks
- Twitter sentiment analysis
- Spam filter using machine learning algorithms
- Proficient in deep learning tools like keras, theano

Figure 3: Portion of a sample machine learning profile

4.1. Tokenization

The sentences in the resume are segmented to obtain tokens. The delimiter for tokenization is the space character. Tokenization involves breaking up of the portion of the machine learning resume into tokens as depicted in Figure 4. The sentence in the resume, "sentiment analysis and classification using tensorflow in MIT", is converted into a list of tokens comprising of individual words "sentiment", "analysis", "and", "classification","using","tensorflow","MIT".

Figure 4: Tokens

4.2. Stop Word Deletion

Usually the resumes of the candidates are filled with redundant words such as 'is, and' etc. Such terms are called stop words. The removal of such simple stop words from the tokens obtained in the previous step resulted in a 28% rise in the efficiency of the classification model. The elimination is important because inclusion of stop words in the training set would result in false learning by the classifier, which would limit the efficiency of classification. The tokens, generated in the last step involve a stop word "and". The word does not hold significance in the classification of the resume. Hence, the word is deleted from the tokens as shown in Figure 5.

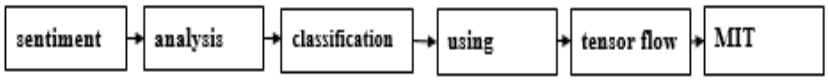


Figure 5: Tokens after stop word deletion

4.3. Parts of Speech Tagging

POS tagging is the next step followed in the NLPP. The step involves tagging the Part of Speeches to each of the tokens obtained after eliminating the stop words. English language has 8 different part of speeches: Verb, Noun, Adjective, Adverb, Pronoun, Preposition, Conjunction, Interjection. The tools and technologies used, Projects undertaken by the candidate is a noun or a pronoun. Since, the allotment of domain to the resume is dependent mainly on these features, only the tokens labelled as nouns and pronouns (NNP) are considered for the next step of NLPP. This step reduces the computation as the number of words considered for classification is reduced greatly. The POS tagging [7] of the tokens is shown in Figure 6. Only the tokens labelled as nouns move to the next stage in the pipeline.

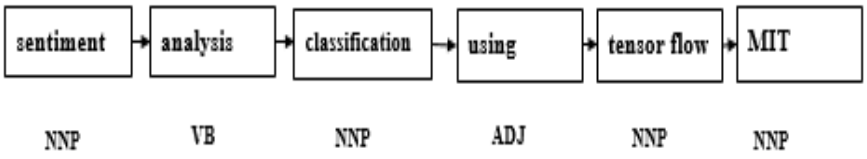


Figure 6: Tokens with POS tagging

NNP: Noun or a pronoun
ADJ: Adjective
VB: Verb

4.4. Named Entity Recognition

All the tokens labelled nouns and pronouns are subjected to Named Entity Recognition [8]. The tokens include the names of the candidate, educational institutions and place names. These tokens are eliminated in the present stage of the pipeline by identifying the candidate names, organizations and place name tokens, which are of trivial importance for allotment of a project to the candidate. This is the final step of the pre-processing pipeline. The tokens emerging out of this step are considered by the classifier for classification. The tokens which are recognized as name, place or organization in this step are eliminated as shown in figure 7. Only the contents shown in Figure 8 move to the classification module.

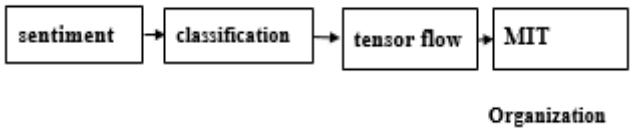


Figure 7: Tokens with Named Entity recognition

The process is repeated for every sentence in the resume portion shown in Figure 3.The final set of tokens are then forwarded to the classifier module as shown in Figure 8**.**

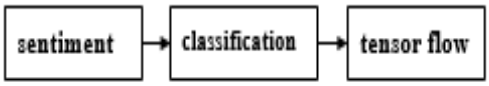


Figure 8: Output from NLPP

5. CLASSIFICATION MODULE

The classifier module receives the set of tokens from the NLPP module and is responsible for the classification of resumes into specific domains. The flowchart depicting the various steps involved in the module is shown in Figure 9. The classifier module is an ensemble learning [14] based voting classifier consisting of 5 different machine learning classifiers. The learning algorithms constituting the voting classifier are Naïve Bayes [6], Linear SVC, Bernoulli NB, Multinomial Naïve Bayes, Stochastic Gradient Descent classifiers respectively.

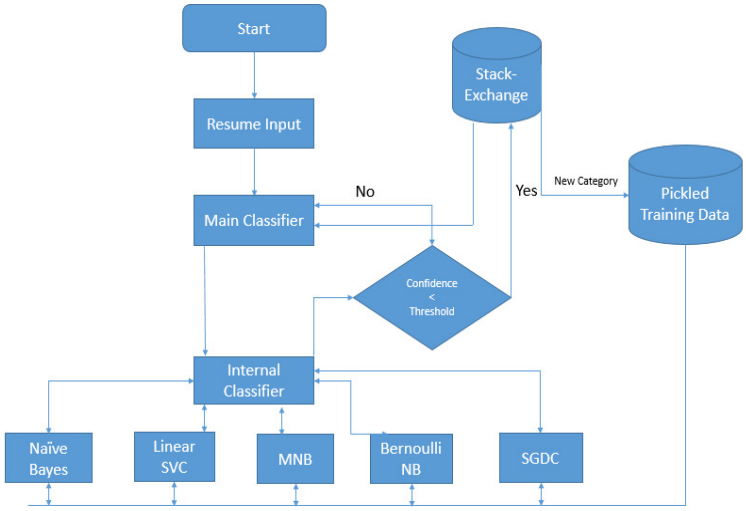


Figure 9: Flow chart of classification module

The set of individual learning classifiers constituting the ensemble learning based voting classifier are commonly used learning algorithms in the field of machine learning and form a benchmark for any new classifier. The initial set of domains identified are mentioned in Table 1. The training data set comprised of 30000 profiles of employee allocated to projects in the domains mentioned in table 1, based on their interests and work experiences previously. The data set formed the basis of the training set (supervised) of the individual machine learning classifiers constituting the ensemble learning based voting classifier. The dataset was divided into training set and a test set in 9:1 ratio. Hence, 27000 profiles formed the training set, while the rest 3000 formed the test set.

Table 1: Initial Domains to be mapped with the profile

Domains
Artificial Intelligence
Computer Architecture
Computer Graphics
Databases
Distributed Computing
Computer Networks
Web Technologies

Each classifier after training predicts the test data based on the learning from the training set. Efficiency of the classifier is the percentage of number of right predictions by the classifier on the test data. The efficiency of each individual classifier is shown in the Table 2.

Table 2: Efficiency of individual classifiers

Name of the Classifier	**Efficiency of prediction in %**
Naïve Bayes	79
Linear SVC	83
MNB	93
Bernoulli NB	89
Logistic Regression	81

The classifiers are allotted votes based on their respective efficiencies as shown in Table 2. This ensures higher influence of the classifier having greater efficiency in categorizing the tokens to a domain than the classifier having less efficiency, while mapping the profile to a domain. The number of votes given to each classifier is governed by equation 1.

$$\text{Number of votes to a classifier X} = \frac{Efficiency\ of\ classifier\ X}{Efficiency\ of\ classifier\ with\ least\ efficiency\ in\ training\ data} \quad (1)$$

From Table 2, the least efficient classifier is Naïve Bayes (79% efficient), hence the denominator for the above equation 1, is 79%. Hence, for Linear SVC (83% efficient), the number of votes allotted is83%/79% = 1.05 votes. The votes allotted to individual classifiers is given in Table 3.

Table 3: Distribution of votes between different classifiers

Name of the Classifier	Number of Votes
Naïve Bayes	1
Linear SVC	1.05
MNB	1.17
Bernoulli NB	1.12
Logistic Regression	1.025

Confidence of the main classifier for the selection of domain for the set of tokens is the ratio of the number of votes casted in favour of the majority class to the total number of votes available with the individual classifiers.

$$\text{Confidence} = \frac{\text{Number of votes cast in favor of majority domain}}{Total\ number\ of\ votes} \quad (2)$$

Total number of votes cast by the classifiers are the sum of the votes allotted to the individual classifiers. From Table 3, the total votes allotted = 1 (Naïve Bayes) + 1.05 (Linear SVC) + 1.17(MNB) + 1.12(Bernoulli NB) + 1.025(Logistic Regression) = 5.365. If Naïve Bayes, Linear SVC and Bernouli NB choose machine learning domain for the set of tokens from Figure 8, while the rest of the classifiers chose distributed computing for the same set of tokens. Machine learning domain has a total vote share of 3.17, while distributed computing has a share of 2.195. Hence, the machine learning domain with a majority vote share of 3.17 is the domain selected by the main classifier. The confidence of the resume classifier from the equation 2 is given as 3.17/5.365 = 0.590.

Higher confidence of the model depicts higher consensus of the classifiers in selecting the domain and higher the chances of the classifier being correct in the decision to classify the set of tokens into the right domain. There are some instances, when the votes of the classifiers are cast equally on number of different categories. In this case, to break the deadlock, the domain which has been voted by the most efficient classifier is chosen.

6. DYNAMIC CLASSIFICATION

A threshold value is set on the confidence with which the resume is classified by the main ensemble learning based voting classifier. The threshold value set is 0.55, which is slightly greater than half of the total votes allotted to the classifiers. When the main classifier in the classification module predicts a set of tokens from the NLPP module with a confidence value less than the threshold value set, the classifier withholds the classification of the instance and calls the REST API trained on stack exchange dump to find a new category for the resume. A confidence value less than 0.55 signifies an absence of consensus among the individual classifiers in allotment of project to the resume.

Figure 10: Portion of sample cloud computing resume

Figure 10 depicts a portion of sample cloud computing profile. Cloud computing does not come under any domain listed in Table 1. The sentence in the resume, "specialized in Microsoft Azure", forms a set of tokens after passing through the NLPP pipeline as shown in Figure 11.

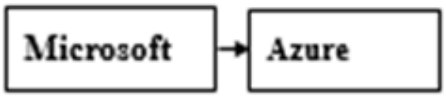


Figure 11) Tokens generated after NLPP

The individual classifiers listed in Table 2 are unable to classify the set of tokens generated by the NLPP pipeline because of the restricted knowledge of only those domains which are listed in Table 1. In such cases, the classifiers do not reach a consensus on the domain to be classified for the given resume. The model withholds the classification, calls the REST API which returns a plethora of related topics. The new topics returned from the API is based on the results from the Association Rule Learning carried out on the dump of Questions and Answers of Stack Overflow. Latent Dirichlet Allocation (LDA)is used for the topic modelling [12]. In natural language processing, LDA is a statistical model that allows sets of observations to be explained by unobserved groups that clarify why some parts of the data are similar.

Once the topic modelling is completed, the new domain is added to the list of domains in Table 1. The new data gathered in the process is added to the training set under the new domain. The internal models in Table 2 are then re-trained on the updated training set and the efficiency of the classifier is recalculated. A new domain of cloud computing is added to the Table 1, after topic modelling of the data returned from the REST API for the tokens shown in Figure 11.

7. RESULTS

Figure 12 depicts the output of the tool for the portion of machine learning resume shown in Figure 3.The model correctly classifies the instance of the resume into artificial intelligence domain. The classification is driven by the identification of deep learning tools like "tensorflow", "keras", "theano" and machine learning buzzwords like "spam filter", "analysis and classification", "fraud detection" in the resume.

Naïve Bayes, Linear SVC and Bernoulli NB classified the resume into AI domain while Multinomial NB and Logistic Regression classified the resume to Distributed computing and Computer Architecture respectively.

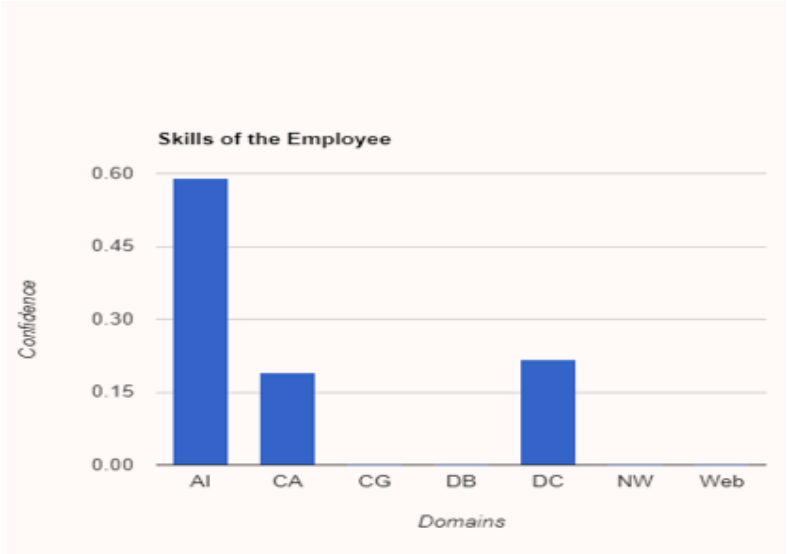


Figure 12: Output graph

AI: Artificial Intelligence
CA: Computer Architecture
CG: Computer Graphics
DB: Databases
DC: Distributed Computing
NW: Networks
Web: Web Technologies

.

The X axis forms the domains in the initial training set, while the Yaxis depicts the confidence of the main classifier across various domains. The profile is then allotted a project in the AI domain as classified by the classifier.

$$\text{Confidence for AI} = \frac{\text{Number of votes for AI}}{Total\ number\ of\ votes}$$

$$= \frac{\text{Votes of Naive Bayes+Linear SVC+ BNB}}{Total\ number\ of\ votes}$$

$$= \frac{1+1.05+1.12}{5.365}$$

$$= 0.590$$

$$\text{Confidence for Distributed Computing} = \frac{\text{Number of votes for Distributed Computing}}{Total\ number\ of\ votes}$$

$$= \frac{\text{Votes of Multinomial NB}}{Total\ number\ of\ votes}$$

$$= \frac{1.17}{5.365}$$

$$= 0.218$$

Confidence for Computer Architecture = $\frac{\text{Number of votes for Computer Architecture}}{Total\ number\ of\ votes}$

= $\frac{\text{Votes of Logistic Regression}}{Total\ number\ of\ votes}$

= $\frac{1.025}{5.365}$

= 0.191

The efficiency of the voting based classifier was 91.2%, predicting domains accurately for 2736 out of the total 3000 resumes in the test set, while in 80% of such cases, the confidence of the model was above 0.7. The accuracy without the re-training was 84.2%. Hence, the re-training after topic modelling of the related topics returned from the REST API increased the efficiency of the classifier by 8.3%.

Increase in efficiency after retraining= $\frac{efficiency\ after\ retraining - efficiency\ before\ retraining}{efficiency\ before\ retraining} * 100$ (3)

= $\frac{91.2-84.2}{84.2} * 100$

= 8.3%

8. CONCLUSION AND FUTURE WORK

The results from the model are encouraging. The promise of a tool which can help the HR team in making better decisions relating to the project allotment of a new recruit is met by the classification model. The ensemble learning based voting classifier performs extremely well compared to the individual classifiers while predicting most of the instances of the test data. This is because of the fact that the confidence of the model while categorizing resumes is influenced by the majority of the votes cast by the individual classifiers rather than a single classifier. An increase in the efficiency of classifier is observed on retraining with the information returned from the association rule mining [11] of stack-overflow questions and answers dataset. The future scope of the project is to build an ensembling deep neural network model [13] in place of the ensemble learning based classifier. Data hungry deep learning models especially Generative Adversarial Networks [15], can harness the huge amounts of unstructured and structured data available in the organisation, thus increasing the efficiency of the resume classification.

ACKNOWLEDGEMENTS

We would like to record our appreciation to all the people involved in the research and development of the resume classifier application. We express our gratitude to the In Step team at Infosys for providing us with the requisite dataset involving the resumes of the employees which formed the training set of the classifier.

REFERENCES

[1] A Comparative Study on Different Types of Approaches to Text Categorization, Pratiksha Y. Pawar and S. H. Gawande International Journal of Machine Learning and Computing, Vol. 2, No. 4, August 2012

[2] Web Document Classification Based on Fuzzy k-NN Algorithm Juan Zhang Yi NiuHuabeiNie Computer and Information Computer and Information Computer and information China.

[3] V Ram, Prasanna, "A unique way of measuring the similarity of the documents using neural networks ,International Journal of Engineering Research and Development, Vol.2, no.6, July 2013, pp 397-401.

[4] Yiteng Huang, Jingdong Chen Blind, "Classifying the text using the power of deep learning", International Journal of Engineering and Technology ,Taipei, Taiwan, 2013, pp 3153-3156.

[5] Daniel R and George V, "A Latent Semantic Analysis method to measure participation quality online forums", 2016 IEEE 16th International Conference on Advanced Learning Technologies, January 2016, pp 108-113.

[6] Jongwoo Kim, Daniel X. Le, and George R. "NaïveBayesClassifier for Extracting Bibliographic Information from Biomedical Online Articles", national Library of Medicine,8600 Rockville Pike, Bethesda, MD 20894, USA

[7] Natural Language Query Processing Using Semantic Grammar international Journal Of Computer Science And Engineering Vol II Issue II March 2010 pg no 219-233

[8] Natural Language Query Processing international Journal of Computer application And Engineering Technology and Science IJ-CA-ETS Oct 2009 pg no. 124-129

[9] https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/

[10] Karin Spark Jones, "A statistical interpretation of term specificity and its application in retrieval", Journal of Documentation, vol. 28, no. 1, pp. 11-21, 1972.

[11] T. Griffiths, M. Steyvers, "Finding scientific topics", Proceedings of the National Academy of Sciences, vol. 101, pp. 5228-5235, 2004.

[12] D. M. Blei, A. Y. Ng, M. I. Jordan, "Latent Dirichlet allocation", J. Mach. Learn. Res., vol. 3, pp. 993-1022, 2003.

[13] Z.-H. Zhou, J. Wu, W. Tang, "Ensembling neural networks: Many could be better than all", Artificial Intelligence, vol. 137, no. 1-2, pp. 239-263, 2002.

[14] Z. Zhu, Q. Chen, Y. Zhao, "Ensemble dictionary learning for saliency detection Image And Vision Computing", Elsevier, vol. 32, pp. 180-188, 2014

[15] Lin Zhu, Yushi Chen, PedramGhamisi, JónBenediktsson, "Generative Adversarial Networks for Hyperspectral Image Classification", Geoscience Remote Sensing IEEE Transactions on, pp. 2018

AUTHORS

Suhas Tangadle Gopalakrishna:

Suhas works at Infosys Limited as a Specialist Programmer as part of Expert Track. His research interests include natural language processing, computer vision, human-computer interaction. He is active in various programming forums including Hacker Rank, Hacker Earth and Stack Overflow.

Dr Vijayaraghavan Varadharajan:

Vijayaraghavan is a Principal Research Scientist at Infosys Limited doing research in the field of data analytics, Searchable encryption, Security assessment, Cloud security, Authentication and Privacy protection. He has over 17+ years of experience in the fields of research, industry and academia. Prior to that he served as an Assistant Professor and guided many post graduate professional students. He has many granted US patents in key technology areas, published many research papers in International journals and conferences and also served as a Technical Reviewer, Program Committee member and Chair for many conferences around the globe.

AUTHOR INDEX

Amir Moradi 37

Ary Setijadi Prihatmanto 63

Carmadi Machbub 63

Dian Andriana 63

Egi Muhammad Idris Hidayat 63

Fangyan Zhang 01

Guo-ming Chen 09

Hong-zhou Tan 09

Kazuki Miyazawa 23

Liang Xue 09

Masahiro Muraguchi 23, 47

Mehran Ektesabi 37

Saman A. Gorji 37

Shun-dao Xie 09

Suchart Yammen 37

Suhas Tangadle Gopalakrishna 77

Sureerat Tang 37

Takuya Kazama 23

Vijayaraghavan Varadharajan 77

V. Mahesh K. Reddy 37

Wen-fang Wu 09

Xiong-yong Zhu 09

Yasuhiro Shimazu 47

Yu Sun 01

Yushi Shirato 47

Zehao Li 01