Technology Forces

Journal of Engineering and Sciences

ISSN 1994-862X (Print)
ISSN 2790-0053 (Electronic)

Volume 4 - Issue 2 | July - December 2022

KIET Research Projects

KIET

Karachi Institute of Economics & Technology

Technology Forces

Journal of Engineering and Sciences

EDITORIAL BOARD

College of Engineering
Karachi Institute of Economics & Technology

ADVISORY BOARD

For queries on submissions, please send an email to tfjes@kiet.edu.pk

COLLEGE OF ENGINEERING

Vision

"To impart highest quality engineering education with an aim to produce proficient, creative and ethical electrical engineers who can take the role of visionary leaders, entrepreneurs and researchers at national and global levels".

Mission

"To educate students through highly qualified faculty to have strong theoretical and practical expertise, team-work, leadership qualities, ethical values and entrepreneurship skills. Establish and maintain state-of-the-art laboratories, latest curriculum and industrial linkages. Prepare graduates to pursue higher studies and undertake research at national and global levels".

AIMS AND SCOPE

Technology Forces Journal is a peer-reviewed journal that aims at the publication and dissemination of original research articles on the latest developments in all fields of engineering, science and technology. The journal publishes original papers in English which contribute to the understanding of engineering science and improvement of the engineering technology and education. Papers may be theoretical (including computational), experimental or both. The contribution should be unpublished before and not under consideration for publication elsewhere. Technology Forces Journal maintains a standard double-blind peer review process. The double-blind process means that the identity of the author and the reviewer are not known to each other.

Focused areas for publication include; but not limited to:

- Signal Processing
- Data Mining and Visual Analytics
- Software Engineering & Future Generation Computing
- Embedded Systems
- Smart Networks
- Molecular and Nanoelectronics
- Applied Controls and Power

OPEN ACCESS POLICY

For the benefit of authors and research community, this journal adopts open access policy, which means that the authors can self-archive their published articles on their own website or their institutional repositories. The readers can download or reuse any article free of charge for research, further study or any other non-profitable academic activity.

PEER REVIEW POLICY

Peer review is the process to uphold the quality and validity of the published articles. Technology Forces Journal uses double-blind peer review policy to ensure only high-quality publications are selected for the journal. Papers are referred to at least two experts as suggested by the editorial board. All publication decisions are made by the journal's Editors-in-Chief on the basis of the referees' reports. We expect our board of reviewing editors and reviewers to treat manuscripts as confidential material. The identities of the authors and reviewers remain confidential throughout the process.

COPYRIGHT

DISCLAIMER

The opinions expressed in **Technology Forces Journal** are those of the authors and contributors, and do not necessarily reflect those of the journal management, advisory board and the editorial board. Papers published in Technology Forces Journal are processed through double blind peer-review by subject specialists and language experts. Neither the **College of Engineering** and nor the editors of **Technology Forces Journal** can be held responsible for errors or any consequences arising from the use of information contained in this journal, instead; errors should be reported directly to the corresponding authors of the articles.

Table of Contents

THE CASE OF ENERGY GENERATION & USAGE IN SPAIN

Muhammad Atif Saeed, Saim Ahmed, Muhammad Ismail Mansoor
Shaheed Zulfiqar Ali Bhutto Institute of Science and Technology, SZABIST, Karachi, Pakistan
Email: m.atif@szabist.edu.pk (Muhammad Atif Saeed), Ismail.mansoor@szabist.edu.pk (Muhammad Ismail Mansoor)

Received: 18-June-2022 / Accepted: 9-December-2022
Karachi Institute of Economics and Technology || Technology Forces Journal, Volume 4, Issue 2, 2022

ABSTRACT

The study is based on an open-access dataset of energy demand and consumption in Spain. This four-year dataset includes data on Spain's electrical consumption, generation, pricing, and weather. Several visualization techniques were used to estimate the pattern that could be used to develop the country's sustainable demand and supply model. According to the analysis, the highest energy consumption hours in Spain over the last four years have been between 1500 and 2000, while energy consumption is lowest between 0000 and 0500 in the morning. Furthermore, the findings revealed that energy consumption is highest in February and lowest in September. As a result of this analysis, the government can now implement a sustainable energy policy that boosts output during peak periods.

Keywords: Energy Forecasts, Sustainable Systems, Data Mining, Efficient Scheme.

1. INTRODUCTION

Energy efficiency can be define as practicing less energy in a more efficient way to do the same amount of work, especially avoiding energy excess. As efficient usage of energy has tremendous benefits ranging from lowering greenhouse gas radiations to reducing demand for power imports while taking down home and economy-wide expenditures. Although alternative energy can help to reach these mentioned objectives, enhancing energy efficiency is the most economical and often fastest way to reduce fossil fuel practice. Every area of economy including metropole transportation, state architectures, big industries or power production have excellent opportunities for efficiency improvements. The four year open access database, 'Electricity demand dataset' includes Spain power generation and consumption with respect to its prices and climate data. ENTSOE, a public gateway for Transmission Service Operator (TSO) data, was used to obtain consumption and generation statistics. The Spanish 'TO Red Electrica Espana' provided the settlement prices. Weather data for the five main cities in Spain was gathered through the Open Weather API as part of a personal project and made accessible online.

The database is special since it's includes hourly data on energy use as well as TSO demand and pricing forecasts. This allows future estimates to be compared to industry-standard forecasts. By visualizing load and marginal supply curves, some of the major factors affecting electricity demand and supply is mentioned below.

- Energy Saving Building Design

A good effective cost effective building design can contribute to 30% global energy savings

[1]. Any commercial or residential building could be design or modify to harness and utilize maximum energy from nature. This can be achieved by following standard architectural steps like painting it with light color (so that it could reflect back sunrays) to keep temperature maintained and reduce the use of air conditioning unit. Construct it sun facing to reduce the usage of artificial lights in daytime. Switching to energy efficient appliances such as DC inverter technology can also help is reducing the load. [2]–[4].

- Minimal Generation and Distribution losses

Another factor accounts for the wastage of energy is line loss. Although, there is enough technology available for power generation sector to switch from old transmission system on to a modern, however; good capital is required. These line losses may also decreased by replacing transformers that are incorrectly sized, using better conductor for power line connections and introducing capacitor banks. [5], [6].

- Effective urbanization

If the city is master planned, then often its commodity and traveling issues are resolved. The better network of transportation help citizens to opt for public transportation instead of travelling in their own. A bus can accommodate 50 people in an area that is much lesser if they use 50 individual transports, hence; resulting much lesser fuel consumption, reduction in global warming. [7], [8].

- Advanced Vehicles

Electric vehicle can contribute to fuel efficiency if the energy generated to charge them comes from renewable energy sources instead of fossil [9], [10].

- Change in Human Behavior

The major contributor in energy savings is human itself. The society needs to be educated to invest in modern technology, to take home appliances that will give those benefits in longer run, do not look for short and rapid solutions instead contribute in this global goal and lead as an example [11], [12].

1.1 Objective

The research objective is to determine the days' time, day and the month when the most power is utilized in order for alternative energy conservation measures can be implemented. The dataset is based on four years of hourly power consumption tracking to discover the characteristics of minimum and maximum energy dissipation seasons.

2. METHODOLOGY

To achieve the goal, numerous data visualization techniques are used on a given dataset to forecast and identify the desired outcome. Fig 1 depicts the procedure we will use throughout this investigation.

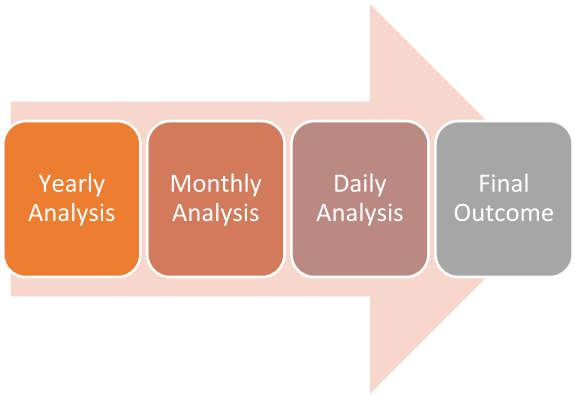


Fig. 1: Process Flowchart

2.1 Yearly Analysis

Data is initially broken down on an annual basis for better visualization, which aids in establishing the hour of the day when the utmost power is used, as perceived from 'Fig. 2'. Since the year's hours are counted in a commutative manner (e.g., one year has 8760 hours), it is possible to establish which quadrant of the year has increased or reduced energy consumption. Likewise, it is important to determine which time of the day consumes more energy; as a result, the mean yearly energy request by each hour of the day is draw against the average power dissipation, as presented in Fig. 3. Furthermore, the most well known statistic in electrical engineering for assessing load usage is the load

distribution plot, which is drawn in between the commutative year's hours versus the average energy usage. The load distribution graph is an extremely valuable module since it may help in determining whether season or month consumes the most or least power, whereas with Fig. 2 we could have done the same thing. The load distribution graph analyzes and clarifies things that are easy to interpret. 'Fig. 4' illustrates the load distribution graph.

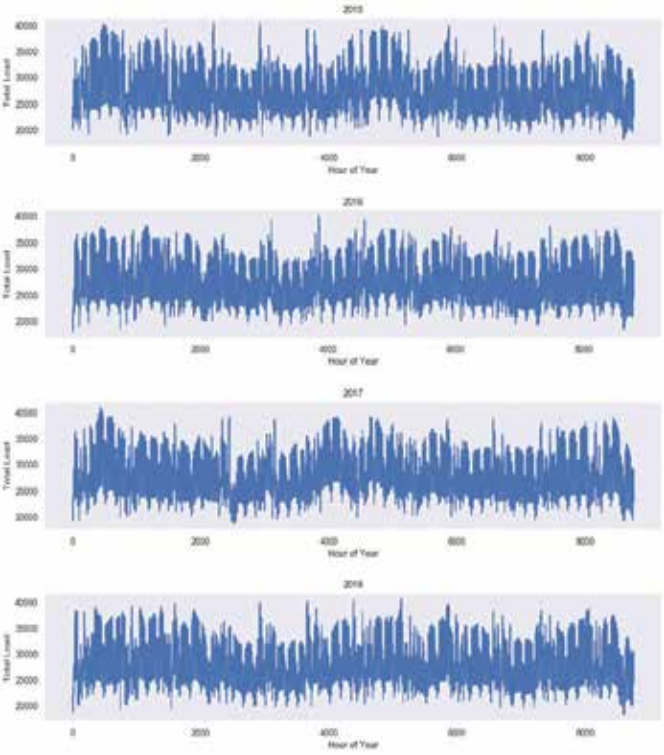


Fig. 2: Year by year energy utilization.

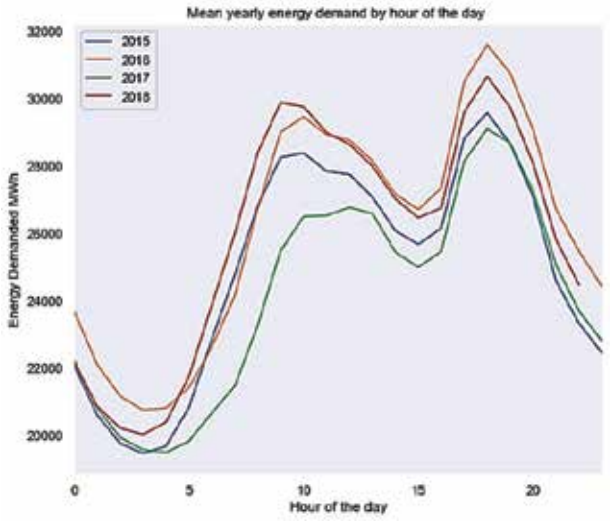


Fig. 3: Mean yearly power requirement by day's hour.

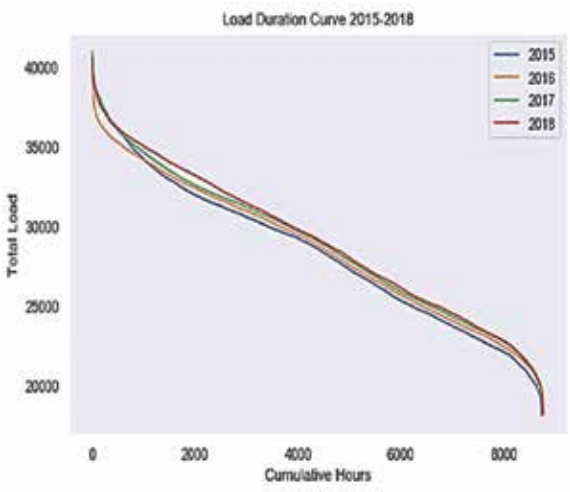
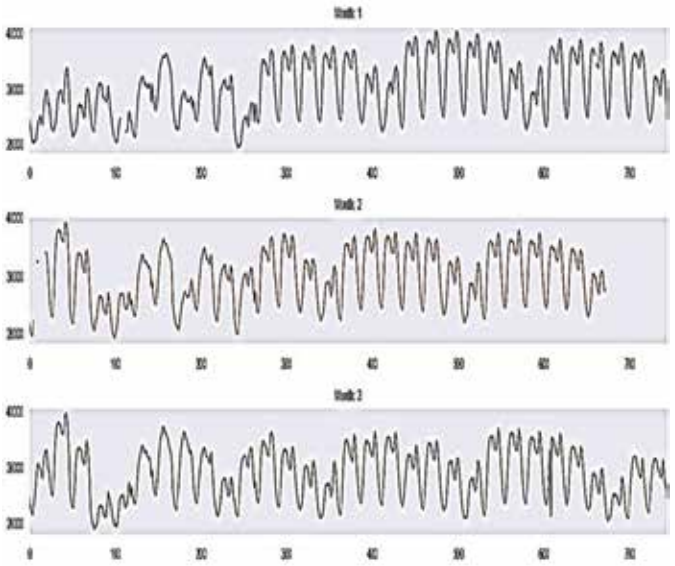


Fig. 4: Load Duration curve

2.2 Monthly Analysis

This part narrows down the year-long investigation for clarity and efficiency. Each month's commutative average consumption is compared to 30 day's hours, i.e. 720 hours. Figure 5 depicts the plots. The outcome will aid in determining the month with the highest electricity consumption. Furthermore, the same data is used to calculate the mean monthly energy usage for each day's hour when compared to the month. Figure 6 depicts the plot. In the same way that the load distribution curve for the year was produced, the load distribution curve for the month was determined in order to determine the optimal month.



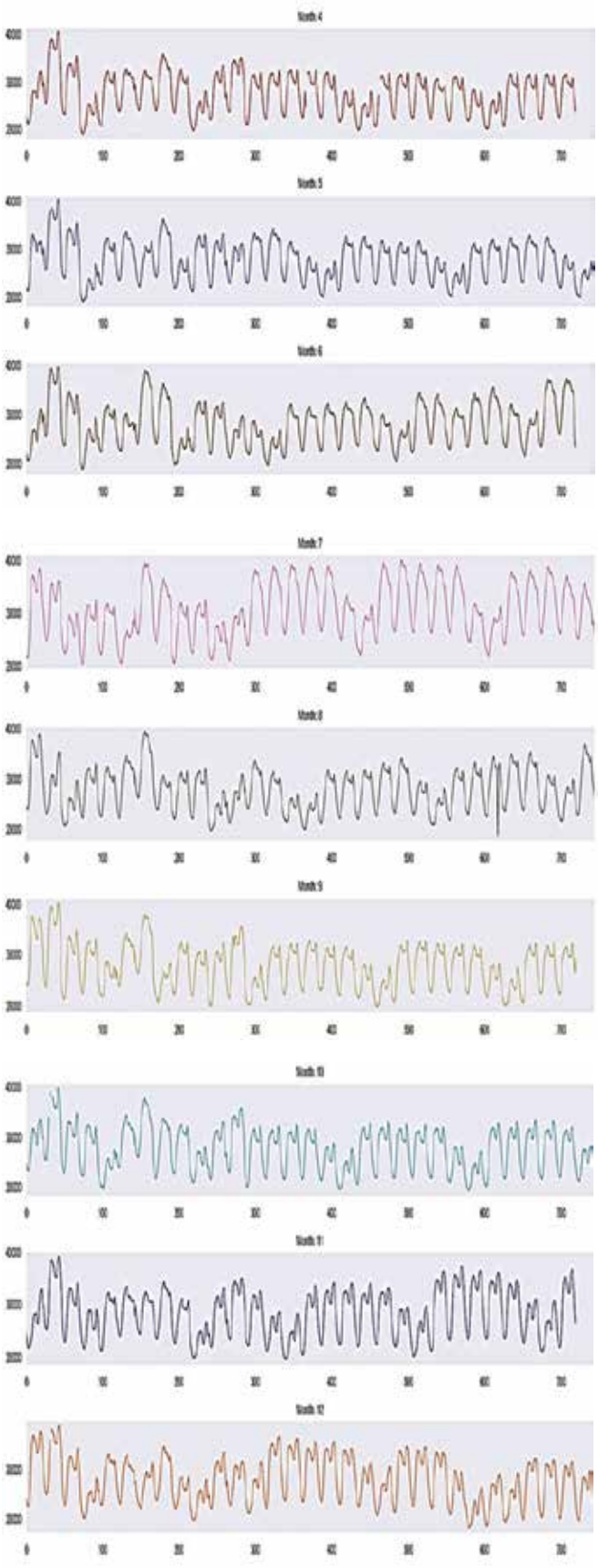


Fig. 5: Monthly average energy breakdown.

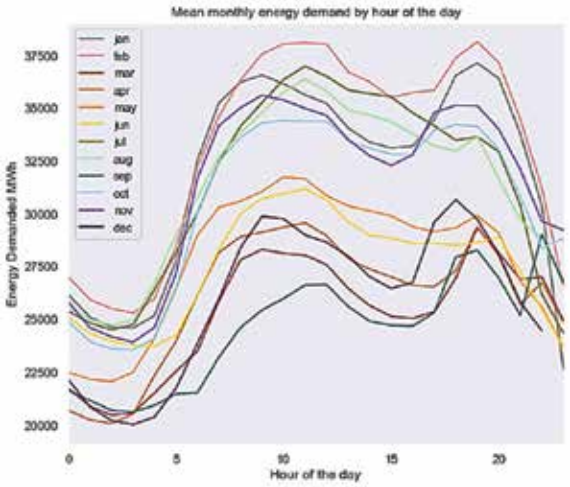


Fig. 6: Mean monthly power utilization by day's hour.

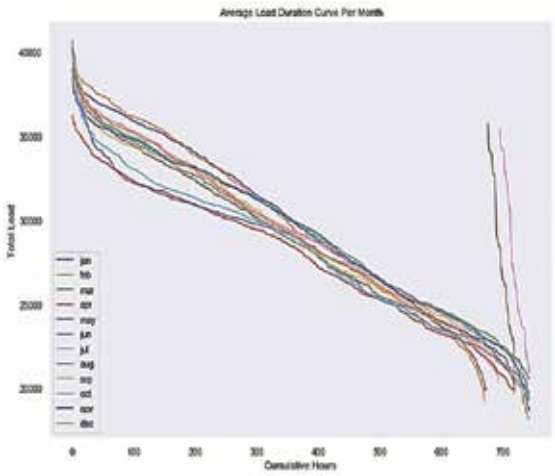


Fig. 7:Average Load Distribution Curve per month.

2.3 Daily Analysis

The daily examination of electrical utilization is shown in Figure 8. This is done to find the average hour of the day when the most energy is consumed.

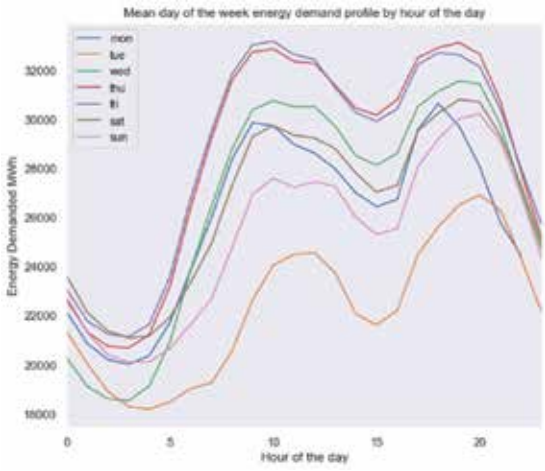


Fig. 8: Energy consumption profile by day's hour on a typical weekday.

3. RESULTS

In this section, the results from the previous section are evaluated and interpreted. As previously said, the research primary goal is to discover the best day's hour, day as well as the season for the minimum and maximum power use. For the previous four years, the peak energy consumption hours have been between 1500 and 2000, whereas the energy usage between 0000 and 0500 hours in the morning has been small. 'Fig. 4' shows load duration graph where the energy utilization is highest at the beginning of the year, that is of January, February, and March; however, the trend abruptly decayed towards the end of the year, showing that energy consumption is lowest in the months of September, October, November, and December.

'Fig. 6' shows that energy usage peaks in February from 1500 to 2000, while it peaks in September from 0000 to 0500 hours. In 'Fig. 7', the load period curve for months revealed that the power utilization trend followed the similar downward trend as the year. Utilization is highest at the month beginning and lowest at the end. Finally, Fig. 8 demonstrated that energy consumption is highest during the week on Mondays from 1500 to 2000, and lowest on Tuesdays from 0000 to 0500 hours.

4. CONCLUSION

According to this study, the most energy is utilised on Mondays between 1500 and 2000 in the month of February, based on the previous four years. On Tuesdays between 0000 and 0500 in September, however, the energy use is modest. As a result of this discovery, the highest energy should be created in February, when consumption is at its peak, and the minimum energy should be generated in September.

REFERENCES

[1] A. M. Al-Ghaili, H. Kasim, N. M. Al-Hada, M. Othman, and M. A. Saleh, "A Review: Buildings Energy Savings - Lighting Systems Performance," IEEE Access, 2020, doi: 10.1109/ACCESS.2020.2989237.

[2] L. Pérez-Lombard, J. Ortiz, and C. Pout, "A review on buildings energy consumption information," Energy Build., 2008, doi: 10.1016/j.enbuild.2007.03.007.

[3] M. Economidou, V. Todeschi, P. Bertoldi, D. D'Agostino, P. Zangheri, and L. Castellazzi, "Review of 50 years of EU energy efficiency policies for buildings," Energy and Buildings. 2020. doi: 10.1016/j.enbuild.2020.110322.

[4] M. Sztubecka, M. Skiba, M. Mrówczynska, and A. Bazan-Krzywoszanska, "An innovative decision support system to improve the energy efficiency of buildings in urban areas," Remote Sens., 2020, doi: 10.3390/rs12020259.

[5] S. Chowdhury, S. P. Chowdhury, and P. Crossley, Microgrids and active distribution networks. 2009. doi: 10.1049/pbrn006e.

[6] H. S. Ahn, B. Y. Kim, Y. H. Lim, B. H. Lee, and K. K. Oh, "Distributed Coordination for Optimal Energy Generation and Distribution in Cyber-Physical Energy Networks," IEEE Trans. Cybern., 2018, doi: 10.1109/TCYB.2017.2669041.

[7] International Energy Agency, "Key World Energy Statistics," Statistics (Ber)., 2011, doi: http://dx.doi.org/10.1787/weo-2011-en.

[8] S. MacDonald, B. Winner, L. Smith, J. Juillerat, and S. Belknap, "Bridging the rural efficiency gap: expanding access to energy efficiency upgrades in remote and high energy cost communities," Energy Effic., 2020, doi: 10.1007/s12053-019-09798-8.

[9] T. Energy, "Annual Energy Review," 2010.

[10] M. Śmieszek, N. Kostian, V. Mateichyk, J. Mościszewski, and L. Tarandushka, "Determination of the model basis for assessing the vehicle energy efficiency in urban traffic," Energies, 2021, doi: 10.3390/en14248538.

[11] U.S. Energy Information Administration (EIA), "International Energy Outlook 2017 Overview," U.S. Energy Inf. Adm., 2017.

[12] I. U. Hussaini and N. H. A. Majid, "Human behaviour in household energy use and the implications of energy efficiency delivery: A case of Bauchi, Nigeria," Int. J. Energy Sect. Manag., 2014, doi: 10.1108/IJESM-10-2013-0005.

HAPTIC SITUATIONAL AWARENESS DEVICE USING CONTINUOUS VIBROTACTILE SENSATIONS

Muhammad Aakash Khaliq, Hammad Munawar
Department of Avionics Engineering, College of Aeronautical Engineering, National University of Sciences and Technology, Islamabad,44000, Pakistan
Email: a.khaliq@cae.nust.edu.pk (Muhammad Aakash Khaliq), h.munawar@cae.nust.edu.pk (Hammad Munawar)

Received: 13-September-2022 / Accepted: 12-December-2022
Karachi Institute of Economics and Technology || Technology Forces Journal, Volume 4, Issue 2, 2022

ABSTRACT

For humans interacting with their environment (such as walking in crowded or low light conditions) or operating vehicles (aircraft or cars), automated alerts and directional cues serve to guide their attention towards the priority task. Traditionally, such alerts and direction cues are provided using audio-visual channels. In fast changing situations, the human audio-visual resources can get degraded. In such situations, vibrotactile alerts and directional cues can easily grab the attention of operators. Furthermore, vibrotactile cues allow for more heads-up time. This research investigates the effects of placing a vibrotactile display on the torso for enhancing situational awareness of humans whose visual/auditory senses may be degraded. The authors describe the findings regarding placement and properties of the proposed device and assess the added effectiveness of the vibrotactile cuing method when used alone and in combination with visual-auditory sensory modalities. In pursuance of this research, a vibrotactile display worn on the torso has been developed, which consists of low-cost vibrotactile actuators that provide vibrotactile alerts and directional cues. Alerts and directional information are encoded and presented using the Funnelling Illusion. Funnelling Illusion provides perceptually more natural cues while reducing the number of actuators, thus enhancing reliability and significant reduction in weight and cost of the device. Experimental results on human volunteers indicate that vibrotactile sensations which are intuitive to perception of the user (continuous sensation instead of discrete) contribute more to enhancement of situational awareness of user during sensory distraction as compared to visual-audio cues only..

Keywords: Vibrotactile; multi-sensory haptic, situational awareness; funneling illusion, psychophysical evaluation.

1. INTRODUCTION

Haptics is the use of sense of touch to convey information to the brain. It is further divided into tactile haptics and kinaesthetic haptics, where tactile haptics deals with the pressure, shear or vibration felt on the skin. When humans interact with their environment, tactile haptic sensations are naturally generated and play a very important part in enhancing their situational awareness of the environment. These sensations can also be created artificially with specialized actuators and are sensed by the human skin through mechanoreceptors.

Mechanoreceptors are further divided into two categories, Meissner corpuscles and Pacinian corpuscles. Meissner corpuscles lie in the glabrous skin (non-hairy skin like the hands) and are sensitive to vibration frequencies ranging from 5 to 50 Hz. Pacinian corpuscles lie in the non-glabrous skin (hairy skin) are sensitive to vibration frequencies ranging from 40 to 400 Hz [1].

When humans are interacting with the environment (such as walking in crowded or low light conditions) or operating vehicles (such as aircraft or cars), automated alerts and directional cues serve to guide their attention towards the task in hand and help them avoid potential dangers. These alerts are usually delivered using audio/visual channels in the form of displays, warning lights, audio commands, beeps, tones etc. However, in certain situations (e.g., during hazardous/emergency situations) when a lot of information is presented through multiple methods, visual and audio sensory channels may degrade. In this condition, an additional channel of information like sense of touch can be effective in providing critical information to brain. Moreover, this method of giving information through vibrotactile sensation is silent and inconspicuous to nearby users, giving it a certain advantage over other methods. This is the motivation for developing a vibrotactile display that assists with situational awareness. This is supported by the *Multiple Resource Theory* [1], which states that humans possess multiple fixed-capacity resources (known as sensory channels or sensory modalities) for processing information obtained from multiple sources (such as audio, visual and touch). Thus, when one resource is not available, information can be provided to the brain using other vacant resources. In this context, the term multi-modal depicts multiple display sources (visual, audio, vibrotactile) used together.

Systems that enable the sense of touch to present information are known as *Tactile Displays* [2],[3],[4],[5],[6]. Tactile displays consist of specialized actuators mounted inside specialized clothing, head gear, gloves, bands, or belts. These actuators enable the information to be presented to the wearer in the form of sensations which may be discrete or continuous. Generation of discrete sensations requires multiple vibrotactile actuators. On the other hand, continuous sensations can be generated using minimum number of actuators by specific control of intensity and frequency of vibration and temporal order. Continuous sensations provide a more natural feel (when they are generated within the peak sensitivity region of mechanoreceptors) [7], [8].

In addition to the above-mentioned sensations, a phenomenon known as *Vibrotactile / Tactile Illusions* can be generated [9]. It is possible to trick the human body to perceive touch sensations that are different than what are being generated. Vibrotactile illusions occur on the skin when two actuators are activated in a coordinated pattern that their sensation is felt as combined instead of separate sensations. There are two types of tactile illusions, *Funneling Illusion* and *Sensory Saltation*. In case of *Funneling Illusion*, if a touch sensation is produced at two locations of a limb in a specified sequence, human will not feel any of these touch sensations, but feel a phantom sensation in between the two. It is also possible to control the exact location of this phantom sensation. Location of the phantom is controlled by varying the intensity and temporal order of vibrations [8]. Similarly, in the *Sensory Saltation / Cutaneous* Rabbit series of short vibration pulses are successively rendered at discrete locations on the skin, and the user feels as though a rabbit is hopping on the skin. In order to successfully generate haptic illusions, precise knowledge of human biology, electronics and programming is required [7]. In literature, multiple actuator technologies have been used in creating vibrotactile displays such as *Linear Electromagnetic* actuators, *Rotary Electromagnetic* actuators, and non-Electromagnetic actuators. C-2 actuator, Haptuator Mark-II, Linear Resonant Actuator (LRAs) are examples of linear electromagnetic actuators, Eccentric Rotating Mass (ERM) motors are examples of rotary electromagnetic actuators and Shape Memory Alloys (SMAs) are examples of non-electromagnetic actuators [10][11]. Details

of the actuators have been discussed in hardware implementation section. These actuators are usually known by different names including tactile actuators, tactors and vibrotactile actuators.

Vibrotactile displays find applications in many areas. Alarms and alerts are commonly being used in mobile phones that present some vibration patterns to draw the attention of the user [2][3]. It is common experience that in a noisy environment like a football stadium, visual and auditory channel of a mobile phone user may be overloaded causing her/him to miss an audio alert. However, the chances of missing a vibration alert are significantly less.

More advanced information such as navigation cues can also be presented through wearable tactile devices that guide the user [5]. Direction information can also be provided to people with visual and hearing disabilities through vibrotactile sensation on their skin [12]. Dogs can also be trained to take vibrotactile commands instead of vocal commands from humans[13]. Vibrotactile devices for navigation and obstacle avoidance of disabled persons sitting on power wheelchair have also been developed [14]. Similarly, vibrotactile cues find applications in challenging environments such as flying an aircraft, where warning messages must be robust and clear so that attention of the pilot could be drawn towards important information. Same is the case for landing of (helicopters) in degraded visual environment [15]. In addition to that, collision avoidance information can also be presented to astronauts when they must perform space walks [4].

In this paper, our research contributes towards developing a low-cost wearable vibrotactile display with commercially available vibrotactile actuators for displaying complex information to human wearers through continuous vibrotactile sensations. These sensations have been generated by the human sensory illusion called *Funneling Illusion* [7]. Proposed device consists of a belt worn on the torso, a control box, flexible polyurethane foam for vibration isolation and vibrotactile actuators, as shown in *Figure 1*.

Figure 1 :Proposed vibrotactile belt worn on the torso

This paper has been arranged in 6 sections. Section 2 comprises of related work from literature. In Section 3 hardware and implementation of the proposed device is discussed in detail. Section 4 is about the human experimentation of the device. Results are described in Section 5 while conclusions are drawn in Section 6.

2. LITERATURE REVIEW

In this section relevant work presented in literature has been reviewed. For comparison purposes, different benchmarks have been selected like source of information to be encoded in terms of vibrotactile patterns, identification of coding parameters (vibration patterns and location of information presented on body) used in the existing devices, vibrotactile actuators used to generate vibration patterns, size of the array / numbers of actuators being used, application area and experimental results.

Comparison of related devices that provide situational awareness information is shown in Table 1. Though devices in literature provide continuous vibrotactile navigation information on legs [16] and hands [17], it is important to note that when user must interact with the environment and perform different tasks, then legs and hands are not static. This would cause hindrance in movement and reduce the effect of the vibrotactile display. Our proposed device provides the novel dimension of taking information from surrounding and guides the user for navigation purposes. Furthermore, as our

device is worn on torso, user can freely interact with the environment without any restriction. Effectiveness of the device may increase because it does not hinder the movement of the user so it may be suitable for the applications where user needs to move hands and legs multiple times like walking, operating machines, flying an aircraft, and driving vehicle.

3. VIBROTACTILE ILLUSIONS

A phenomenon called Sensory Saltation was discovered back in 1970 at the Princeton Cutaneous Communication Laboratory while the word *saltation* means "jumping" in *Latin*. Initially three tactors were placed in line on the forearm of a person where the first one was placed close to the wrist. Then three very short pulses were given to the tactile actuators (tactors). Rather than just feeling the vibratory sensation on different locations of the forearm (locations where the tactors were placed), the subject felt as if a rabbit is hopping on the skin from the wrist towards the elbow. Hence, the phenomenon was named *Cutaneous Rabbit*.

After the discovery of this phenomenon, extensive research has been done on cutaneous rabbit type sensations at Princeton University and it was observed by experiments that in order to create a rabbit like sensation the inter-tactor distance should not be greater than 10 cm [18]. Later on, Tan et.al [19] studied the influence of vibrotactile illusions like *sensory saltation* and *cutaneous rabbit* and developed a tactile navigation system with a 3-by-3 tactor array worn on the back. For inter stimulus interval, a great effort has been made in a study by Chouvardas et.al [20], where a judgement is made on the quality of continuous movement that *duration of stimulus (DOS)* is important in creating tactile illusions.

Research continued and multiple types of illusions were discovered. They are now termed as *Vibrotactile Illusions / Tactile Illusions*, most common amongst which are the *Funnelling Illusion* and *Sensory Saltation*. Funnelling Illusion happens on the skin when two actuators are activated in such a way that a virtual phantom appears on the skin midway between the actuators. Location of this phantom depends upon the relative vibration intensities of surrounding actuators and this location can be controlled by varying the vibration intensity of the actuators (which were used to create the phantom). When the intensity or temporal order is varied, virtual phantom feels like moving from one actuator towards the other. This is a well researched tactile illusion [7],[8],[21],[22],[23], [24],[25].

Moreover, modification in the Funnelling sensation can be done by two methods: *Temporal Inhibition and Amplitude Inhibition* [26],[27]. Temporal inhibition phenomenon happens when two discrete vibro-tactile actuators are activated with same intensities but at different time intervals. In result of this activity, funnelled sensation is perceived to be moved toward the earlier activated actuator. This phenomenon has been shown in the Figure 2.

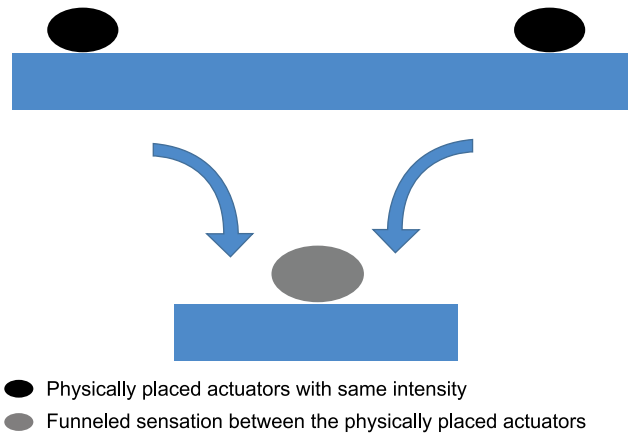


Figure 2: Funneling Illusion because of Temporal Inhibition

In case of amplitude inhibition, vibrotactile actuators are activated with varying intensities but at the same time intervals. In result to this activity, funneled stimulus is perceived to be located near the actuator activated with higher intensity. This phenomenon has been shown in the Figure 3.

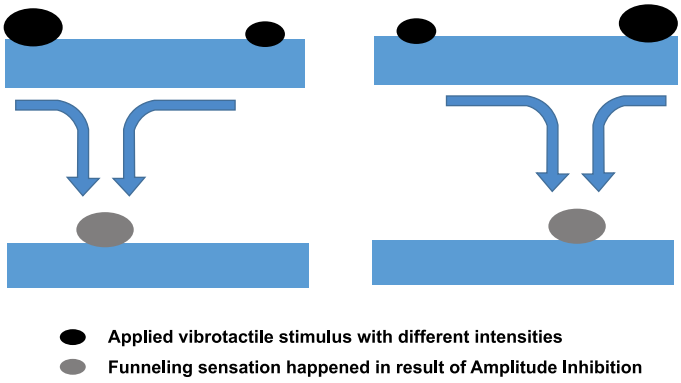


Figure 3: Funneling Illusion because of Amplitude Inhibition

The difference between Sensory Saltation and Funneling Illusion is such that in the former approach user feels discrete sensation like something is crawling on the skin. However, user feels discrete taps instead of continuous sensation [19]. On the other hand, in case of Funneling illusion, by using amplitude and temporal inhibition, continuous sensations and variety of haptic / tactile effects are possible to create with good control over the parameters. With temporal inhibition virtual phantom appears on the skin between the physically placed vibrotactile actuators. Similarly, with the help of amplitude inhibition, generation of continuously moving sensation is also possible [21].

In this paper, we have implemented the hybrid approach of using the amplitude and temporal inhibition. We activated the actuators at varying time intervals (a property of temporal inhibition) and used varying intensity of stimulus (a property of amplitude inhibition). By systematically applying the above-mentioned technique, we can generate continuous moving sensation.

4. PROPOSED DEVICE (HARDWARE DESIGN AND IMPLEMENTATION)

In this section hardware design and implementation of the proposed device has been discussed. A vibrotactile display system consists of *Vibrotactile Actuators, Actuator Array, Vibration Pattern Generation, and Information Encoding*. Therefore, each of sub systems for the proposed device are discussed in the following paragraphs.

A. Vibrotactile Actuators

Vibrotactile sensation is artificially generated through vibrotactile actuators. Multiple types of actuators are available for this purpose. Some of them are custom designed for more reliability and robustness for specific application domains and others are more general in their use. Selection of a suitable vibrotactile actuator is important decision in designing a vibrotactile system. Two main hardware decisions are taken by the designer: firstly, what type of actuator to use and secondly how to arrange them spatially so that vibrotactile sensation can be adequately felt by the user [10].

Examples of vibrotactile actuators available are Eccentric Rotating Mass (coin type), Linear Resonant Actuator (LRA), Eccentric Mass Rubber (EMR), Haptuator Mark 2 and C3 tactors. Shape of the above-mentioned actuators has been shown in the Figure 4 and comparison based on features is shown in Table 2. In the proposed device, ERM is used for its affordability, availability, mechanical simplicity, electrical simplicity, moderate customizability and moderate expressiveness [28], [10].

Table 1: Comparison of most relevant situational awareness devices

Sr. No	Device 1	Device 2	Device 3	Device 4	Device 5	Device 6	Device 7
Author	Devigne *etal.* [14]	Liao *etal.* [16]	Rahal *etal.* [21]	Hoffmann *etal.* [30]	Khaled etal. [28]	Ouyang *etal* [29]	Rupert *etal.* [6]
Developed by	University College London (UCL)	Tohoku University, Japan	University of Ottawa	University of Iceland	Ecole Miitaire Polytachnique, Algeria	Southeast University, China	US ARMY / NAVY
Tactile Actuators	Pico-Vibe	Coin Type Vibration Motor	Pancake DC vibration motor	ERM	(ERM)	DC motor	C-3, EMR
Tactile Array	4-4	6	4	16	16	20	16
Form Factor	Armband	Elastic Band	Armband	Belt	Vest	Vest	Belt
Location	Arm	Lower Leg	Arm	Lower back	Torso	Torso	Torso
Vibration Pattern	Phantom Sensation	Phantom Tactile Sensation	Linear and Logarithmic Intensity variation	Apparent shift of direction	Not made public	Intensity, rhythm, modes of vibrations	Not made public
Encoded Information	Navigation of power wheel chair	Collision alerts	Continuous motion on skin	Illusory upward or downward indication	Flight path climb, Turn rate	Roll & pitch	Hovering of helicopter
Experimentation	Laboratory Environment	Laboratory Environment	Intensity and DOS varied to test the quality of sensation	Users indicated the direction of illusory stimulus	Flight simulator	Flight simulator	T-34 Jet aircraft, UH-60 Blackhawk
Results	Guided the user for obstacles	Perception of vibration decreased	Intensity variation and DOS has significant effect	Activation of actuators aided in creating illusion	Visual + tactile better than visual only	95 % accuracy of rectangular flight patterns	Three times better with TSAS

ERM (model 310-118) by Precision Microdrives has a body diameter of 10 mm, operating voltage of 3-V, optimum vibration frequency of 240 Hz and a weight of 0.8 g that makes it suitable to be used in light wearable devices. ERM is also cost effective and readily available in the market. These features make ERM the best candidate in comparison with other costly and complex actuators. In our design, we have used a frequency range of 200 to 250 Hz [20], [31] as it is best perceivable frequency range of mechanoreceptors in the torso region [31], [32].

1. Biological Background of Human Vibrotactile Sensation

According to biological background, vibration sensing capabilities of humans vary on different body sites and skin types [33]. Four sensory channel theory that identifies afferent neurons to perceive skin deformation, namely *Merkel Disks, Meissner Corpuscles, Pacinian Corpuscles and Ruffini Endings*, is considered as the base of vibrotactile perception. It asserts that: glabrous skin of humans carry all four types of mechanoreceptors [32], [33]. These four types of sensory channels respond to different set of mechanical stimuli that is summarized in the Table 3. In non-glabrous skin Pacinian Corpuscles (PCs) and Rapidly Adapting (RA) mechanoreceptors are the reason behind vibration sensing while Meissner corpuscles are completely absent [9], [10], [34].

2. Location of Stimulus

From biological background of vibrotactile sensation it is evident that vibrotactile system designers should identify the suitable skin type and body location where vibrotactile stimulus is to be rendered. Skin type varies with body location so does the vibration sensing capability of the skin. Human hands and feet have glabrous skin which is most sensitive to even smaller frequency of vibration.

Location of vibrotactile stimulus is chosen according to application. Since our potential application is to provide navigation and directional information to pilots, visually impaired people, and vehicle drivers so out of all body parts we have considered torso to be the suitable one. There are multiple advantages of using torso as location of vibration stimulus, because torso offers huge haptic space (which may be useful for increasing number of actuators in future) and it is mostly static (at least in the case of driving, piloting aircraft and for ground navigation of visually impaired people). Moreover, stimulus provided on the torso does not hinder movement of legs and hands which are already busy in doing other activities. So, torso is the best candidate for being used as a location for vibrotactile stimulus. Furthermore, feasibility of selection of torso has also been discussed in literature [31], [33], [36], [37], [38].

B. Array Design

Proposed design is focused on presenting vibrotactile information on the torso as the torso offers a large haptic space and is mostly static during walking, driving, flying aircraft, and operating machinery. Another key factor to consider while designing an array of vibrotactile actuators is the selection of a suitable vest, belt or elastic strap that could comfortably fit over the body part of the user.

Moreover, robustness of the belt during vigorous movements of the wearer and skin actuator contact is also important. If skin-actuator contact is loose, then vibration stimulation to the skin might be affected. Keeping in view the design requirements, a belt is selected for wearing around the torso. It is made up of nylon webbing for flexibility and better wear ability to any waist size. A square region of 170 by 170 mm is selected on the middle of the belt and an array of five actuators is drawn on the selected region with inter actuator distance of 70 mm [19], [37], [39]. To mount the actuators on the belt, foam of a specific grade and properties is selected to impede the vibrations from spreading elsewhere on the belt. Actuator array design has been shown in Figure 5.

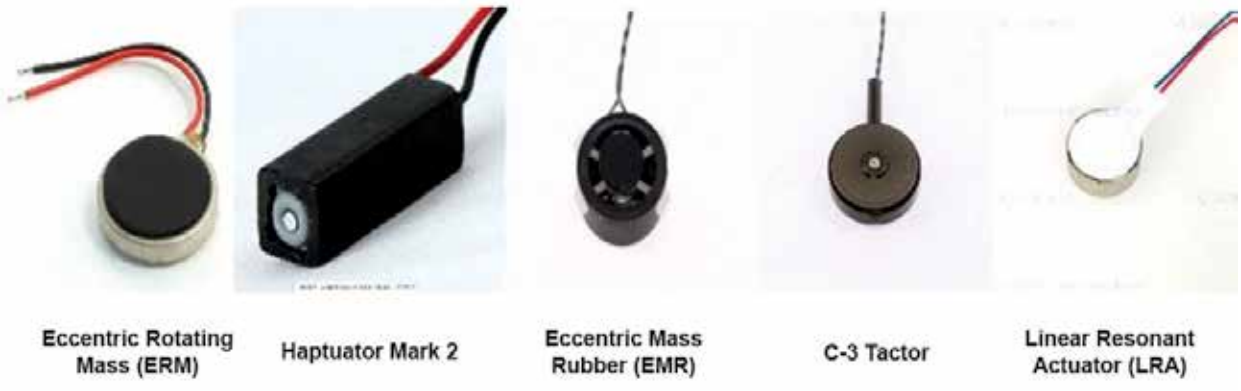


Figure 4:Five commonly used vibrotactile actuators

Type of Actuator	ERM	Haptuator Mark 2	EMR	C-3	LRA
Company	Precision Microdrives	Tactile Labs	Engineering Acoustics	Engineering Acoustics	Precision Microdrives
Weight	0.8 g	9.5 g	5 g	8 g	2 g
Dimensions	10 x 2.1 mm	9 x 9 x 32mm	25.4 x 10.2	20.3 x 6.35 mm	10 x 3.7 mm
Optimum Freq.	240 Hz	120 Hz	115 Hz	240 Hz	175 Hz
Operating Freq.	160 – 280 Hz	90 – 1000 Hz	80 – 140 Hz	180 – 320 Hz	N/A
Cost	6 $	200 $	100 $	300 $	9 $
Availability	High	Low	Low	Low	Medium

Table 2: Features of Vibrotactile Actuators

Characteristics	Merkel Disks	Meissner Corpuscle	Pacinian Corpuscle	Ruffini Endings
Sensory Response	Slow Adapting	Rapid Adapting	Slow Adapting	Rapid Adapting
Skin Depth	Closer to skin surface	Closer to skin surface	Deep beneath surface	Deep beneath surface
Frequency Range (Hz)	<5	3 – 100	10 – 500	15 – 400
Perceived Stimulus	Pressure	Flutter	Skin stretch	Vibration
Shape	Disk	Flattened stack of cells	Branched fibers in a capsule shape	Layered capsule shaped
Location	Between dermis and epidermis	Just below epidermis	Dermis	Deep in the skin

Table 3: Characteristics of Mechanoreceptors in Human skin [9], [10]

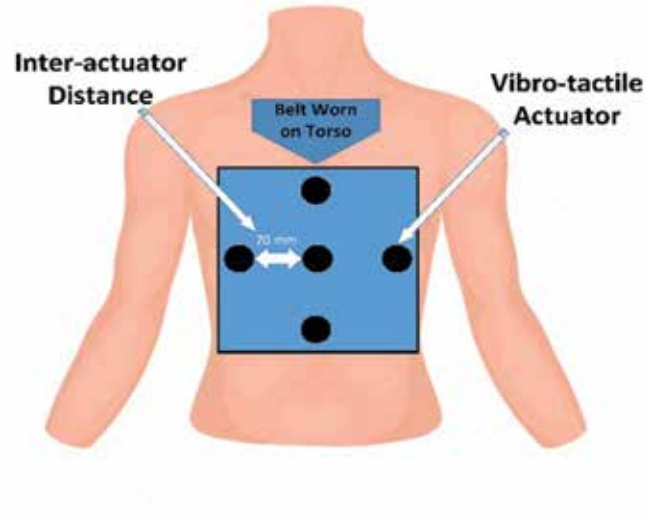


Figure 5: Proposed array of 05 actuators having an inter-actuator distance of 70 mm

Whereas actual design created on belt is shown in Figure 6.

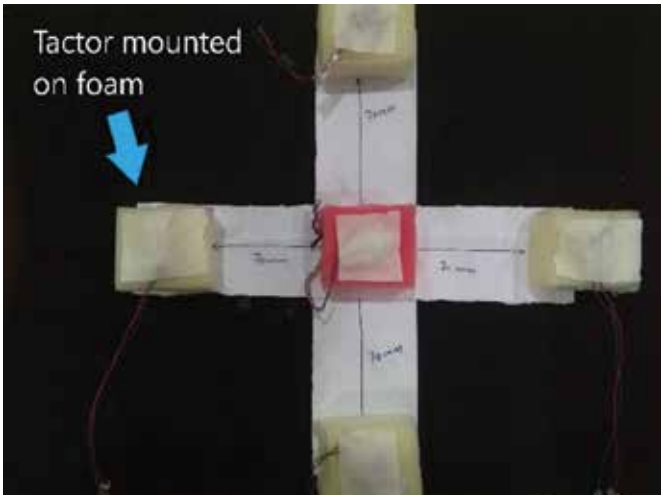


Figure 6: Array of proposed vibrotactile belt

C. Vibration Patterns

For generation of vibration patterns with minimum delay a Control Box has been designed which consists of a micro controller, power supply and necessary circuitry. The circuitry ensures that sufficient starting current is available to vibrotactile actuators so that crisp of haptic effects can be generated. Each sub system (Control Box, Data Communication Interface, and Power Source and Generation of tactile Illusions and continuous sensation) and its operation has been discussed in the paragraphs below.

1. Control Box

The system is designed using an Arduino (NANO) open-source hardware which is based on the Atmel Atmega 128P. It is the main controller that can be programmed through a PC. To address the problem of delays and generation of haptic waveforms, Immersion Corporation's Haptic Libraries were utilized [40]. Haptic driver (DRV-2605) made by Adafruit and originated by Texas Instruments has been used with the Arduino. Haptic driver 2605 forms the core of actuator module and controller [41], [42]. It supports both ERMs and LRAs having four libraries for ERMs and one library for LRAs. Furthermore, it provides flexible control of ERM and LRA over a shared I^2C data bus and relieves the host processor from generating precise pulse-width modulated (PWM) signals. With the help of haptic driver, all the libraries can be accessed and used for different types of vibration waveforms.

Haptic driver is interfaced with micro-controller through data and clock pins. Micro-controller communicates with the haptic driver through I^2C bus and sends data and clock signals. During design phase, configurability of the device has also been considered in terms of change in vibration patterns and Duration of Stimulus (DOS).

2. Data Communication Interface

A very important requirement of the design of the proposed vibrotactile device is its reliability. Taking this requirement in to consideration, all the vibration motors have separate haptic controllers and communication between haptic controllers. All the communication happens through I^2C protocol.

3. Power Source

Portability of the device is important design consideration. For this, a light weight, reasonable backup time and compact sized alkaline battery has been found suitable for the proposed vibrotactile device. Moreover, small size of alkaline battery provides a good charge density as compared to its size. Battery is rated at 9-V and provides almost 1.5-hour backup time for

the continuous operation. Furthermore, it can be operated in a temperature range of -[18]° C~to~[55]° C [43].

4. Important Physical Parameters for Generating Tactile Illusions:

here are certain parameters that should be considered while creating vibrotactile sensations with vibrotactile actuators. For generating tactile illusions their importance further increases, because range of variation of these parameters to be precisely controlled to achieve the best perceivability of vibrotactile sensations. This range depends on the body location that has been selected for presenting tactile illusion cues. These parameters are shown below and each of them are discussed in the following paragraphs.

- Effect of Amplitude
- Frequency of Vibration
- Relation between Speed and Amplitude
- Relation Between Speed and Frequency
- Speed Variations
- Amplitude and frequency prediction difficulties

a) Effect of Amplitude

Practically, human perceive the amplitude of vibration through our skin differently than that measured in controlled laboratory environment. So, vibration sensation felt through our skin may vary due to multiple factors like orientation of the motor, age of the user, location of the device on human body part (hand, arm, torso, and leg), frequency of vibration and other environmental factors. By changing the amplitude of vibration, different haptic effects can be created, and this useful property of amplitude variation is used to create different vibration patterns. Silent messages are conveyed to the user through various vibration patterns where each vibration pattern corresponds to a certain message or information.

b) Frequency

It is important to mention the relationship between amplitude and frequency of vibration. In the case of ERMs, amplitude and frequency of vibration are inseparable. Amplitude of the vibration is magnitude of the vibration waveform generated by ERM while frequency is the period of the vibration waveform. So, in the waveform as magnitude of the vibration increases, frequency of the waveform also increases. It is the difficulty that one quantity cannot be changed without changing the other so most of the time, frequency is altered to alter the amplitude. Since, we know that mechanoreceptors (Pacinian Corpuscle) in skin can perceive vibration frequencies ranging between 40 to 400 Hz while the peak sensitivity of vibration frequency occurs between 200 to 300 Hz [20], [31], [44], [45].

Moreover, another reason to use the peak sensitivity range of frequencies is due to the potential usage of the proposed device in automotive and aircraft environments. Automotive and aircraft environment have many sources of vibration in them like vibrations produced through running engine, air conditioner and road friction, that can also produce vibration at a certain range of frequencies. So, it is best to produce haptic effects at more distinguishable range of frequencies (200 to 300 Hz).

c) Relation of Speed and Amplitude

In any rotational system the relationship between speed and amplitude is exponential instead of linear. Centripetal force (F) produced by off centered mass in ERM can be described by the relation $f=mrw^2$. Here m is the mass of eccentric load, r is eccentricity and w is rotational speed in radians/sec. A minor increase in motor speed will have a greater impact on the amplitude so we can make significant changes in the amplitude by just increasing motor speed.

d) Relation Between Speed and Frequency

Speed and frequency of ERM are one and the same thing. Motor speed is revolutions per minute and vibration frequency is represented in Hz. To convert from Hz to RPM we simply multiply by 60 and to convert from RPM to Hz we divide by 60.

e) Speed Variations

To change the speed of vibration motor we need to change the driving voltage, as driving voltage has direct effect on motor speed. An Increase in driving voltage will increase the torque produced by the motor as the load is fixed motor speed will increase. For reliable motor operation and to get the good lifespan from motor, it is often recommended by the manufactures to drive between two parameters. One parameter is maximum start voltage and other is maximum operating voltage. So, we need to stay within the constraints of voltage, because exposure to higher voltage for a long time may draw current beyond max rated current leading to motor failure.

f) Amplitude and Frequency Prediction Difficulties

It is difficult to predict the real vibration amplitude based on change in voltage because there is not a linear relationship between voltage and amplitude, and it varies with different motors. So, we used the vibration motor performance graph to predict the vibration amplitude and frequency with respect to change in drive voltage that has been specified in the datasheet [46].

5. Creating Continuous Movement Sensation

We want to exploit the human sensory illusion called "*Funneling Effect*" by which we will create continuous motion sensation on human skin. When the vibration sensation is rendered on the human skin, he / she may feel as if something is moving on his / her skin [21]. For this purpose, we control the intensity of the adjacent actuators. We have two actuators separated by a distance of 70 mm which is the proposed inter-actuator distance [39], [47], [48].

In this paper, hybrid approach has been implemented using amplitude and temporal inhibition. Actuators have been activated at different time intervals (a property of temporal inhibition) and used varying intensity of stimulus (a property of amplitude inhibition). By systematically applying the above-mentioned technique, we have been able to generate continuous moving sensation. Inverse proportional intensity technique of amplitude inhibition has been applied in such a way that when intensity of vibration of first actuator increases from low to high value then the intensity of second actuator will decrease from high to low intensity level.

Therefore, the perceived continuous motion will move from the first stimulus position to the second stimulus position. At the same time temporal order of activation of actuators is also controlled in a specified way. This discrete intensity stimulus rendering on different body locations has been felt as continuously moving from one point to the other. By psychophysical influence of the Funneling illusion, we successfully implemented the continuous tactile motion on the skin and have used this continuous tactile motion for presenting directional information.

Intensity variation technique is based on existing research, that by varying the intensity linearly or in a logarithmic way the stimulus can be felt between the actual stimulus locations. The work of [49], [50] describes the linear variation in stimulus amplitudes which causes the sensation to end near the mid-point of both the stimuli.

D. Encoding of Information

Based on Funneling illusion (hybrid approach), direction information has been encoded in the device. So, device can give continuous directional vibration cues. Continuous sensation happening downward indicates "*Forward Movement*" as shown in Figure 7. Vibrotactile sensation occurring from bottom to top (upward) is aimed to command the user for "Backward Movement" as shown in Figure 8. In addition to that, continuous sensation rendered toward right side of the user indicates "*Right Turn*" as shown in Figure 9. Similarly, vibrotactile sensation rendered toward left side of the user signals the user to take "*Left Turn*" as shown in Figure 10. Moreover, vibro-tactile cue for "Stop" command is rendered in such a way that stimuli from four actuators approach the middle actuator as shown in Figure 11. Each directional vibration cue has been rendered on human skin with minimum

of three actuators. Only three actuators have been used for direction information. For stop command, four actuators were activated.

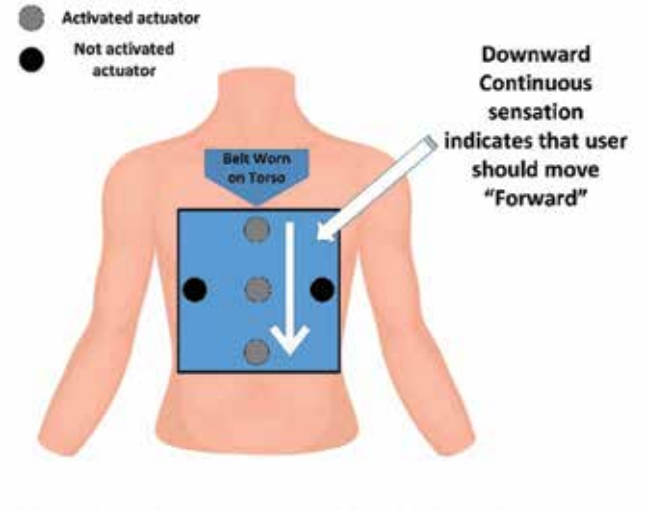


Figure 7: Continuous sensation from top to bottom for indication of "Forward Movement"

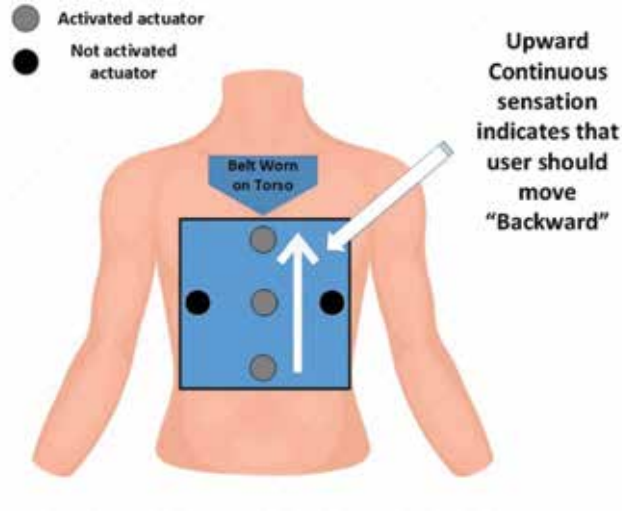


Figure 8: Continuous sensation occurring from bottom actuator toward top actuator signals "Backward Movement"

I. Experimentation

To evaluate the designed vibrotactile display, two important experiments have been performed. First experiment is based on the quality and perception of continuous vibratory sensations created through Funneling Illusion while the second experiment has been performed to evaluate the device performance for ground navigation in a visual / auditory degraded environment.

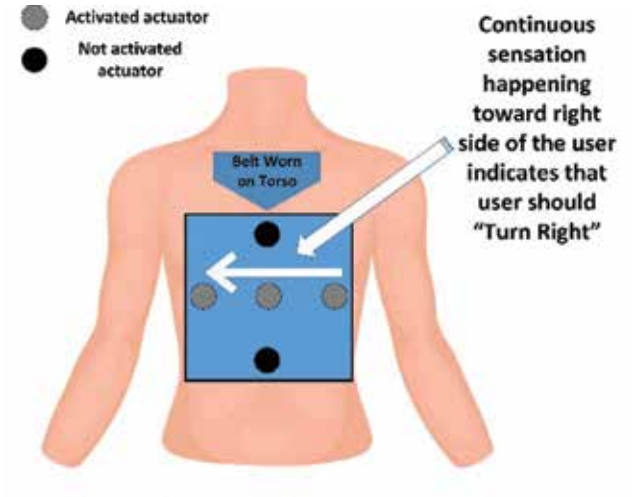


Figure 9: Continuous sensation toward the right side of the user commands the user to "Turn Right"

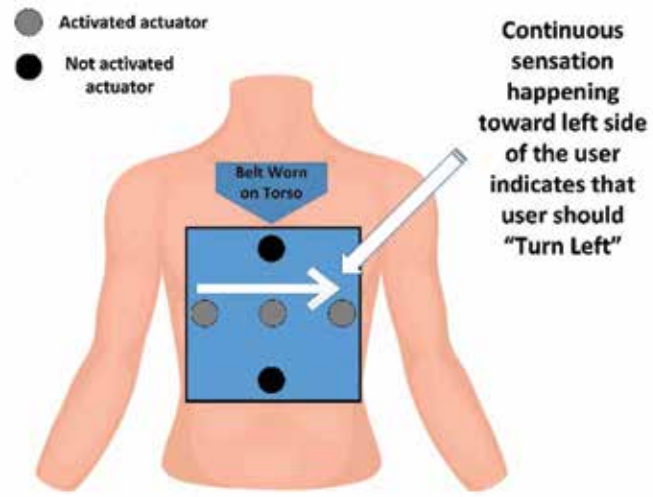


Figure 10: Continuous sensation stimuli happening toward left side of the user signals to "Turn Left"

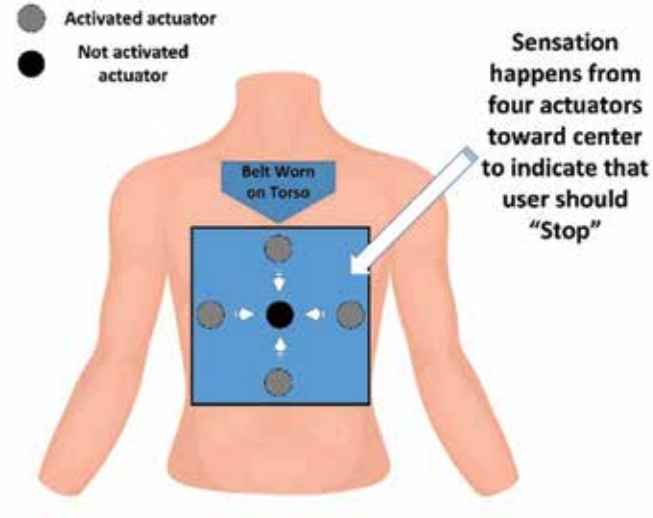


Figure 11: Stimuli approaching from four actuators toward the centre actuator indicate the "Stop" signal

A. EXPERIMENT 1: VIBROTACTILE SENSATIONS

1. Protocol

15 healthy adults (10 males, 5 females) between the ages of 21 to 31 have been selected. Written consent was obtained from the participants. Participants were ostracized from the experiment if they reported any physical injuries or cognitive impairments. All the subjects were naive and did not have any previous knowledge of the vibrotactile device and its working. The proposed device worn on human torso is shown in the Figure 12.

2. Validation of Vibrotactile Sensation

In this scenario the core objective was to assess the quality and perception of continuous sensation on skin based on the feedback provided by the participants. In proposed device, speed of the continuous sensations has been kept programmable and termed as the "*Motion Speed*".

It is directly related to the duration of stimulus of the whole continuous motion in each direction. Motion Speed has two levels based on the duration of stimulus (Fast - 450 ms and slow - 900 ms). We ran both levels (fast, slow) of motion speed parameter for each of the participant and took feedback about it. Out of 15 participants (10 males, 5 females) only two males and one female wanted the sensation to be 'fast' while rest of the participants were comfortable with "slow" level. During this experiment, we only wanted the participants to be comfortable with the vibrotactile sensory modality and be accustomed to it. When the vibrotactile cues were validated successfully, we proceeded toward real-world application (ground navigation) of the device.

B. EXPERIMENT 2: GROUND NAVIGATION

1. Protocol

Each sensory channel is considered as sensory modality. If two sensory channels are involved to present some information they are termed as bi-modal and for more than two sensory channels, a term widely used in the literature is multi-modal (It is the situation where most of the sensory degradation / distraction occurs).

Figure 12: Proposed vibrotactile belt worn by a volunteer

In this experiment each participant was presented with all the seven sensory modality combinations.

- Vibrotactile only
- Visual only
- Audio only
- Visual + vibrotactile
- Visual + audio
- Vibrotactile + audio
- Vibrotactile + audio + visual

Vibrotactile cues were delivered through the proposed device, audio commands were given through headphones while visual cues were obtained by users through their eyes. These cues were enabled/disabled depending on the modalities being presented. For sensory modalities that did not require visual cues, participants were asked to wear black glasses or blindfolds depending on their preference.

Experiment consists of three modes: Single modality, Bi-modal and Multi-modal. First mode to be simulated in the experiment is single sensory modality. As the name suggests, only

one sensory modality has been used in this scenario. Vibrotactile only cues were presented to the participants with device, noise was played in headphones, and they were blind folded to restrict from visual cues. Similarly, when participant was using visual only cues and audio only cues, vibrotactile device was unplugged from the power source.

Protocol for single sensory modality has been shown in Table 4. In the bi-modal scenario, two sensory modalities were used at a time for ground navigation and third sensory modality was removed. Protocol for combination of bi-modal display system has been shown in Table 4.

In the multi-modal display system, all the three sensory modalities (visual, audio, vibrotactile) were used at a time by the participants. Combinations of multi-modal displays have been shown in Table 4. For the participants to navigate in a certain area, a navigation maze was created in a laboratory environment. Maze had several turns that user must cross to reach the destination. A participant wearing vibrotactile belt in the navigation maze has been shown in the Figure 13.

During the experiment, data was taken about reaction time (time taken by the participant to perceive the certain directional cue) and task completion (number of turns / check points crossed through the navigation maze) and saved for use in Statistical Analysis based on Three-Way ANOVA (analysis of variance). ANOVA helped in determining whether the different groups' mean scores differ from one another or not [49], [50].

There are some assumptions that should be taken care of while performing ANOVA. First assumption states that dependent variable must be measured at a continuous level. In ANOVA, second assumption states that each independent variable should consist of at least two categorical groups [51], [52].

In our case, we have designed four categorical groups including gender (two groups: male or female), display conditions or sensory modalities (seven: visual only, vibrotactile only, audio only, visual + audio, visual + vibrotactile, vibrotactile + audio and visual + audio + vibrotactile) and reaction time (in seconds: 1, 2, 3 and 4). Third assumption states that all the data should be collected randomly and independently. Fourth assumption asserts the importance of no significant outliers in the data. The problem with outliers is that they might have a negative impact on three way ANOVA [49]–[52]. All these assumptions were duly considered during our analysis.

Figure 13: Participant wearing vibrotactile belt in navigation maze scenario while goal of the experiment is to guide the participant to walk through the maze from starting point till end point by taking vibrotactile cues from belt

2. Validation of Directional Cues in Ground Navigation

In this experiment, participants were assigned a task to cross the navigation maze from starting point to end point. User must cross six check points / tasks to reach the end point of Navigation maze. Each of this check point / task carries a weight which is "1" for successful check point / task and "0" for unsuccessful check point / task. Completion of all the check points / tasks rewards the participant 6 marks. All the three modes of sensory modalities were used by the participants to perform the navigation maze experiment. In addition to that, in multi-modal combination, participants were intentionally distracted in such a way that additional task of playing game on cell phone and listening to music along with the primary task of crossing navigation maze had been assigned to them. Reason behind creating distraction was to assess the performance of vibrotactile rendering when visual and auditory sensory channels were degraded / distracted.

Furthermore, data of reaction time (time taken by the participant to perceive the direction) and task completion details of all the participants for each of sensory modality were noted for further use in statistical analysis. A participant attending to visual display of cell phone screen, taking audio commands, and receiving vibrotactile navigation cues has been shown in the Figure 14.

I. Results

ANOVA analysis showed some significant main effects for the dependent variable of task completion (Table 5). Moreover, independent variable sensory modality and reaction time has statistical significance, but independent variable gender is statistically insignificant as shown in (Table 5).

A. Effect of Sensory Modality on Task Completion

Two hypotheses: Null hypothesis and Alternate hypothesis have been formulated to assess the effect of sensory modality on task completion. Null hypothesis states that, there is no significant effect of sensory modality on task completion if, the condition ($\rho>0.05$) is true for 95% Confidence Interval (CI). On the other hand, Alternate hypothesis states that there is significant effect of sensory modality on task completion, if $\rho<0.05$ for 95% Confidence Interval (CI).

Figure 14: Participant attending to visual display of cell-phone screen, taking the audio commands / white noise from headphones, and taking vibrotactile cues from belt

Trials depicted significant main effect for the type of sensory modality (Table 5) chosen to perform the navigation scenario, most prominently, when sensory degradation / distraction scenario was performed. Positive effects have been observed when there was distraction / degradation in visual and audio channel but vibrotactile cues helped in task completion.

Moreover, reaction time of the participant was faster when vibrotactile cues were involved as compared to when vibrotactile cues were absent. Sensory modality successfully passed the significance level test. Our Null Hypothesis (there is no significant effect of sensory modality on task completion) is completely wrong since $\rho<0.000$. So, our Alternate Hypothesis (There is a significant effect of sensory modality on task completion) is true. This significance level has been tested for $\alpha<0.05$ and we achieved a 95 % confidence Interval (CI) for the effect of sensory

modality on task completion.

Figure 15 shows the basic differences among the sensory modalities (visual, vibrotactile, audio) and their combinations as shown in Table 4. For task completion and reaction time, sensory display involving vibrotactile cues produced better means than the sensory modalities that does not incorporate vibrotactile cues. Since the experiment in which all the sensory modalities (vibrotactile + visual + audio) have been used, depicts the highest mean for task completion. So, it is evident that sensory modality used for task completion has significant main effect.

B. Effect of Reaction Time on Task Completion

Two hypotheses (Null and Alternate) have been formulated to assess whether there is an effect of Reaction Time on task completion or not. Null Hypothesis states that: There is no significant effect of reaction time on task completion if ($p>0.05$) for 95 % Confidence Interval (CI).

Alternate Hypothesis states that: There is significant effect of reaction time on task completion if $p<0.05$ for 95 % Confidence Interval (CI). Three-way ANOVA was performed, and results show that $p<0.000$ in Table 5 which means that null hypothesis is completely wrong and alternate hypothesis is true. So, our claim of alternate hypothesis is verified, which states that there is significant effect of reaction time on task completion. The significance level / confidence interval (CI) was $\alpha<0.05$.

Reaction time had significant main effect on the level of success of task completion. It is evident from the results in the Figure 15 that participants showed a very quick reaction to multi-modal display when vibrotactile cues were involved while performing the tasks. In multi-modal display, if one channel had been distracted or degraded the rest of the channels were there to provide the useful information to the participant about task completion. Especially the vibrotactile channel relieved the other distracted / degraded sensory channels. It has been observed that when visual and audio channels were distracted / degraded vibrotactile cues helped the user in completion of the task and improved the reaction time of participant for a certain navigation cue.

Moreover, vibrotactile only and combination (vibrotactile + visual + audio) gave almost similar results and their mean scores lie approximately in the same region. So, we can say that when sensory channels are degraded / distracted or not present, vibrotactile only cues can help the user in performing the navigation task with promising results and without distracting / degrading the other sensory channels. Moreover, it is evident from the result shown in Figure 16 that fast reaction time is associated with a greater number of tasks completed and slow reaction time is associated with a smaller number of completed tasks or a smaller number of checkpoints crossed.

Table 4: Combinations of sensory modalities used during ground navigation experiment

Seven Display Conditions	Modes	Vibrotactile Device	Visual Display	Audio Commands
Vibrotactile cues only	Single Modality	ON	OFF	OFF
Visual cues only	Single Modality	OFF	ON	OFF
Audio cues only	Single Modality	OFF	OFF	ON
Vibrotactile + Visual cues	Bimodal	ON	ON	OFF
Vibrotactile + Audio cues	Bimodal	ON	OFF	ON
Visual + Audio	Bimodal	OFF	ON	ON
Vibrotactile + Audio+ Visual cues	Multi-modal	ON	ON	ON

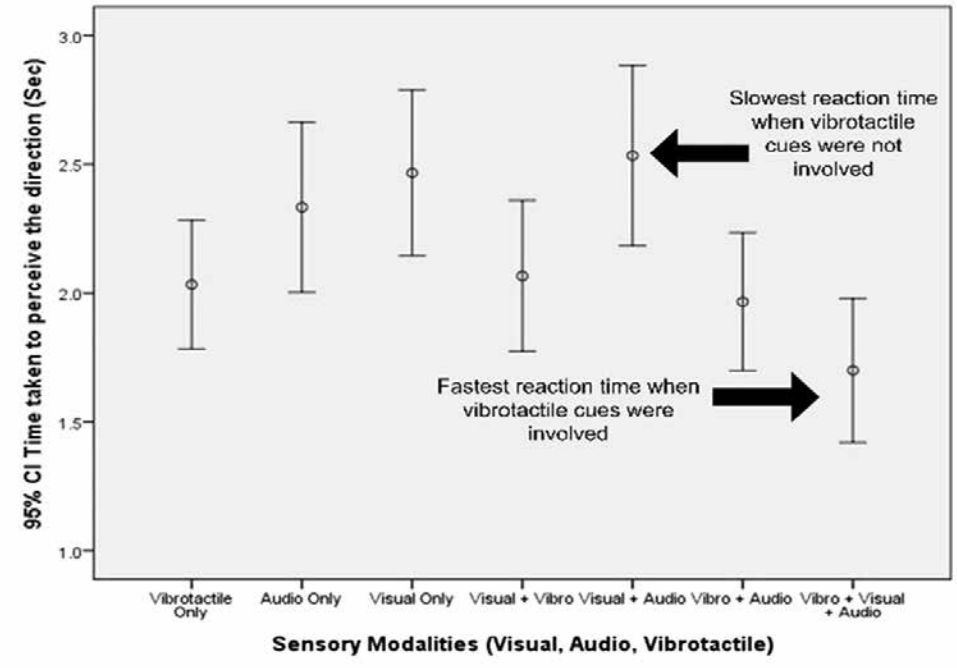


Figure 15: Reaction time (time taken to perceive the direction) of the participant varied with the change of Sensory Modalities

Main Effect	F	d.f	ρ
Sensory Modality	7.770	6	0.000
Reaction Time	172.729	3	0.000
Gender	0.651	1	0.421

Table 5: Significant main effects from three-way ANOVA

C. Effect of Gender on Task Completion

Two hypotheses (Null and Alternate) have been formulated for the main effect of gender to assess whether it has significant effect on task completion or not. Null hypothesis states that: Gender (male, female) does not have any effect on task completion. Null hypothesis is true if $p > 0.05$ for CI $\alpha<0.05$.

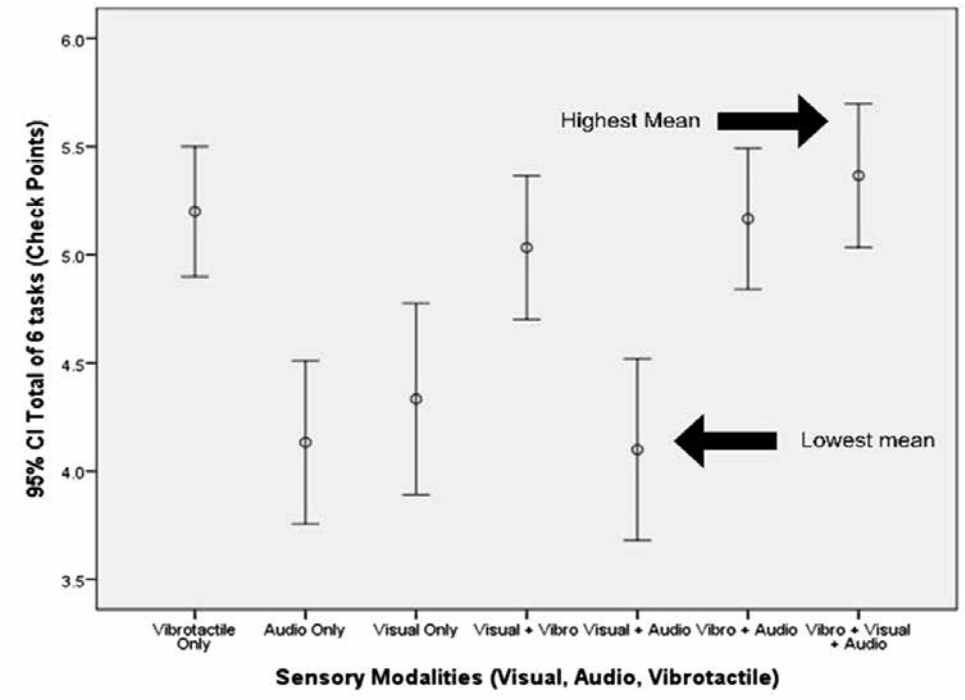


Figure 16: Task completion (check points crossed) has been improved when vibrotactile cues were used independently or in combination with other sensory modalities

While alternate hypothesis states that gender has significant effect on task completion if ($p<0.05$) for Confidence Interval (CI) $\alpha<0.05$. From the results of three-way ANOVA in Table 5 it has been found out that Gender (male, female) does not have significant effect on task completion so our Null Hypothesis is true because ($p > 0.05$). In light of the results, it can be said that gender is not the decisive factor for task completion and device is equally suitable for both the genders.

II. CONCLUSIONS

In this study, the effects of combinations sensory modalities (Table 4) on reaction time and gender then in turn on task completion have been carefully observed. To assess the effect of sensory modality, reaction time and gender on task completion hypotheses have been formulated which state that: change in sensory modality, reaction time and gender will produce significant effect on number of task completion of the participant during ground navigation scenario.

However, with experimental data it is observed that number of task completion has been increased when vibrotactile cues were involved no matter alone or in combination with other displays (visual and audio). Data also depicts the fact that fast reaction time has significant effect on task completion and helped in task completion. Moreover, gender does not play a significant role in task completion as both genders showed almost similar mean of number of task completion and almost equal sensitivity to reaction time.

From results, it is evident that number of tasks completed has highest means for multi-modal display (visual + audio + vibrotactile) while the means for vibrotactile only cues also fall between the mean of multi-modal display. Similarly, mean of vibrotactile + audio cues were not significantly different from vibrotactile only cues.

Results depicted that participant showed fastest reaction time for multi-modal display (visual + audio + vibrotactile) and slowest for bimodal (visual + audio) displays. In addition to that, audio only and visual only means also fall in between and show the similar significance level as bi-modal (visual + audio) feedback. During experimentation, participants reported that they took less time to perceive the correct direction that's why their reaction time was faster when vibrotactile cues were combined with visual and audio cues. This also illustrates that reaction time and task completion performance was better when vibrotactile display was involved.

The non-significant effect of gender is also a finding that gender whether male or female does not show a promising difference on task performance and results were approximately in accordance. So, our dependent variable, number of tasks completed (check points crossed successfully) is not influenced by gender.

III. KEY POINTS

- Vibrotactile situational awareness device has been developed that provides navigation information using continuous vibrotactile sensations.
- Funneling illusion is found to be the most suitable for generating continuous vibration on human skin with minimum number of actuators.
- Results showed that vibrotactile illusions produce continuous sensations and enhance the situational awareness with minimum training time.
- Vibrotactile device allows for more heads-up time and participants completed more tasks (check points crossed) when vibrotactile cues are involved. In future, design ergonomics of the device will be improved.

REFERENCES

[1] C. D. Wickens, "Multiple resources and mental workload," Hum. Factors, vol. 50, no. 3, pp. 449–455, 2008.

[2] C. G. Henshue, G. L. Henshue, and J. C. Rice, "Tactile warning panel apparatus and system with smart technology." Google Patents, 2017.

[3] M. A. Cuddihy, M. K. Rao, and J. Le, "Haptic vehicle alert based on wearable device." Google Patents, 2018.

[4] A. Gibson, A. Webb, and L. Stirling, "Analysis of a wearable, multi-modal information presentation device for obstacle avoidance," in 2017 IEEE Aerospace Conference, 2017, pp. 1–9.

[5] F. Kiss, R. Boldt, B. Pfleging, and S. Schneegass, "Navigation systems for motorcyclists: exploring wearable tactile feedback for route guidance in the real world," in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 2018, pp. 1–7.

[6] B. D. Lawson and A. H. Rupert, "Coalition warfare program tactile situation awareness system for aviation applications: System development," 2014.

[7] Y. Jin, H. Chun, G.-H. Yang, and S. Kang, "Vibrotactile cues using tactile illusion for motion guidance," in 2013 13th International Conference on Control, Automation and Systems (ICCAS 2013), 2013, pp. 1064–1067.

[8] J. Salazar, K. Okabe, and Y. Hirata, "Path-Following Guidance Using Phantom Sensation Based Vibrotactile Cues Around the Wrist," IEEE Robot. Autom. Lett., vol. 3, no. 3, pp. 2485–2492, 2018.

[9] H. Culbertson, S. B. Schorr, and A. M. Okamura, "Haptics: The present and future of artificial touch sensation," Annu. Rev. Control. Robot. Auton. Syst., vol. 1, pp. 385–409, 2018.

[10] S. Choi and K. J. Kuchenbecker, "Vibrotactile display: Perception, technology, and applications," Proc. IEEE, vol. 101, no. 9, pp. 2093–2104, 2012.

[11] L. Jiang and J. M. Cruz-Hernandez, "Haptic device with linear resonant actuator." Google Patents, 2016.

[12] K. Aqeel et al., "Skin stroking haptic feedback glove for assisting blinds in navigation," in 2017 IEEE International Conference on Robotics and Biomimetics (ROBIO), 2017, pp. 177–182.

[13] Y. Golan, B. Serota, A. Shapiro, O. Shriki, and I. Nisky, "A Vibrotactile Vest for Remote Human-Dog Communication," in 2019 IEEE World Haptics Conference (WHC), 2019, pp. 556–561.

[14] L. Devigne et al., "Power wheelchair navigation assistance using wearable vibrotactile haptics," IEEE Trans. Haptics, vol. 13, no. 1, pp. 52–58, 2020.

[15] A. H. Rupert, B. D. Lawson, and J. E. Basso, "Tactile situation awareness system: recent developments for aviation," in Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2016, vol. 60, no. 1, pp. 722–726.

[16] Z. Liao, J. V. S. Luces, and Y. Hirata, "Human Navigation Using Phantom Tactile Sensation Based Vibrotactile Feedback," IEEE Robot. Autom. Lett., vol. 5, no. 4, pp. 5732–5739, 2020.

[17] F. Chinello, C. Pacchierotti, J. Bimbo, N. G. Tsagarakis, and D. Prattichizzo, "Design and evaluation of a wearable skin stretch device for haptic guidance," IEEE Robot. Autom. Lett., vol. 3, no. 1, pp. 524–531, 2017.

[18] F. A. Geldard, Sensory saltation: Metastability in the perceptual world. Lawrence Erlbaum, 1975.

[19] H. Tan, R. Gray, J. J. Young, and R. Taylor, "A haptic back display for attentional and directional cueing," 2003.

[20] V. G. Chouvardas, A. N. Miliou, and M. K. Hatalis, "Tactile display applications: A state of the art survey," in Proceedings of the 2nd Balkan Conference in Informatics, 2005, pp. 290–303.

[21] L. Rahal, J. Cha, and A. El Saddik, "Continuous tactile perception for vibrotactile displays," in 2009 IEEE International Workshop on Robotic and Sensors Environments, 2009, pp. 86–91.

[22] G.-H. Yang, M. Jin, Y. Jin, and S. Kang, "T-mobile: Vibrotactile display pad with spatial and directional information for hand-held device," in 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2010, pp. 5245–5250.

[23] J. Salazar, Y. Hirata, and K. Kosuge, "Motion guidance using haptic feedback based on vibrotactile Illusions," in 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2016, pp. 4685–4691.

[24] K. Zhao, F. Rayar, M. Serrano, B. Oriola, and C. Jouffrais, "Vibrotactile cues for conveying directional information during blind exploration of digital graphics," in Proceedings of the 31st Conference on l'Interaction Homme-Machine, 2019, pp. 1–12.

[25] D. Pittera, D. Ablart, and M. Obrist, "Creating an illusion of movement between the hands using mid-air touch," IEEE Trans. Haptics, vol. 12, no. 4, pp. 615–623, 2019.

[26] D. S. Alles, "Information transmission by phantom sensations," IEEE Trans. man-machine Syst., vol. 11, no. 1, pp. 85–91, 1970.

[27] L. Rahal, J. Cha, A. El Saddik, J. Kammerl, and E. Steinbach, "Investigating the influence of temporal intensity changes on apparent movement phenomenon," in 2009 IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurements Systems, 2009, pp. 310–313.

[28] K. Fellah and M. Guiatni, "Tactile Display Design for Flight Envelope Protection and Situational Awareness," IEEE Trans. Haptics, vol. 12, no. 1, pp. 87–98, 2018.

[29] Q. Ouyang, J. Wu, and M. Wu, "Vibrotactile display of flight attitude with combination of multiple coding parameters," Appl. Sci., vol. 7, no. 12, p. 1291, 2017.

[30] R. Hoffmann, M. A. B. Brinkhuis, Á. Kristjánsson, and R. Unnthorsson, "Introducing a new haptic illusion to increase the perceived resolution of tactile displays," in 2nd international Conference on Computer-Human-Interaction Research and Applications, CHIRA. In review, submitted, 2018, vol. 2, p. 2018.

[31] E. E. L. Gunther, "Skinscape: A tool for composition in the tactile modality," Massachusetts Institute of Technology, 2001.

[32] A. H. Rupert, "Tactile situation awareness system: proprioceptive prostheses for sensory deficiencies.," Aviat. Space. Environ. Med., vol. 71, no. 9 Suppl, pp. A92--9, 2000.

[33] A. Wilska, "On the vibrational sensitivity in different regions of the body surface," Acta Physiol. Scand., vol. 31, no. 2–3, pp. 285–289, 1954.

[34] G. A. Gescheider, J. H. Wright, and R. T. Verrillo, Information-processing channels in the tactile sensory system: A psychophysical and physiological analysis. Psychology press, 2010.

[35] K. O. Johnson, T. Yoshioka, and F. Vega--Bermudez, "Tactile functions of mechanoreceptive afferents innervating the hand," J. Clin. Neurophysiol., vol. 17, no. 6, pp. 539–558, 2000.

[36] K. Schultz, B. McGrath, B. Cheung, G. Craig, S. Jennings, and A. Rupert, "In-flight evaluation of Tactile Situational Awareness System during high hover and simulated shipboard landing," in AIAA Guidance, Navigation, and Control Conference, 2009, p. 6122.

[37] Ó. I. Jóhannesson, R. Hoffmann, V. V. Valgeirsdóttir, R. Unn\thórsson, A. Moldoveanu, and Á. Kristjánsson, "Relative vibrotactile spatial acuity of the torso," Exp. brain Res., vol. 235, no. 11, pp. 3505–3515, 2017.

[38] L. A. Jones and S. J. Lederman, Human hand function. Oxford University Press, 2006.

[39] L. A. Jones, B. Lockyer, and E. Piateski, "Tactile display and vibrotactile pattern recognition on the torso," Adv. Robot., vol. 20, no. 12, pp. 1359–1374, 2006.

[40] Immersion-Corporation, "Haptic Libraries." 2019.

[41] Texas-Instrument, "Haptic Driver 2605." 2020.

[42] Adafruit, "Haptic Driver 2605." 2020.

[43] Energizer, "9-Volt Battery." 2020.

[44] E. GUNTHER, "Composing for the Sense of Touch," in Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02), 2002, pp. 1–6.

[45] S. Brewster and L. M. Brown, "Tactons: structured tactile messages for non-visual information display," in Proceedings of the fifth conference on Australasian user interface-Volume 28, 2004, pp. 15–23.

[46] PrecisionMicrodrives, "ERM 310-118." May 2020.

[47] J. Rosenthal, N. Edwards, D. Villanueva, S. Krishna, T. McDaniel, and S. Panchanathan, "Design, implementation, and case study of a pragmatic vibrotactile belt," IEEE Trans. Instrum. Meas., vol. 60, no. 1, pp. 114–125, 2010.

[48] K. O. Sofia and L. Jones, "Mechanical and psychophysical studies of surface wave propagation during vibrotactile stimulation," IEEE Trans. Haptics, vol. 6, no. 3, pp. 320–329, 2013.

[49] A. Rutherford, Introducing ANOVA and ANCOVA: a GLM approach. Sage, 2001.

[50] K. D. Bird, Analysis of variance via confidence intervals. Sage, 2004.

[51] L. Statistics, "Three-Way-ANOVA." 2020.

[52] R. G. Miller Jr, Beyond ANOVA: basics of applied statistics. CRC press, 1997.

EFFICIENT DATA COMPRESSING TECHNIQUE FOR WIRELESS NETWORKS

Authors: Muhammad Hassan, Dr. Bilal Muhammad Khan*
Department of Electrical Engineering, Faculty of Engineering & Technology, National University of Science & Technology, Karachi, Pakistan
*Correspondence Email: bmkhan@pnec.nust.edu.pk

Received: 20-October-2022 / Accepted: 12-December-2022
Karachi Institute of Economics and Technology || Technology Forces Journal, Volume 4, Issue 2, 2022

ABSTRACT

Data compression, also known as source coding, is a way that can be defined as not to reduce data quality but to make it shorter in quantity. In short, it is a way of expressing a message with fewer bits that can also be easily recovered or reconstruct on receiving end with no ambiguity. There are two ways of data compression, lossless compression and lossy compression with their several sub-categories like Huffman Coding, Arithmetic Coding, Run Length Encoder, Lempel–Ziv Welch (LZW), Burrows Wheeler transform, Move to Front transform and so many more which are going to be discussed in much detail in this paper. Each has its own specific advantages and disadvantages based on which each has its own suitability in various applications.

Keywords: Data compression, lossless compression, lossy compression, Huffman Coding, Arithmetic Coding, Run Length Encoder, Lempel–Ziv Welch (LZW), Burrows Wheeler Transform, Move-to-Front Transform.

1. INTRODUCTION

In modern era, data compression is gaining so much importance as reducing to redundant information continuously repeating in data like blank spaces in graphic and text, same colors in an object or picture reduces the electronic space for data storage and makes its communication faster on a channel. A successful compression is one with which data can also be retrieved in its original form. Today such data compressing technologies have been developed from which 1:10 to 1:20 compression ratio can be achieved. Fax machine was the first device on which data compression was implemented. There are two types of data compression describing below:

1.1. Lossy Compression

Lossy compression is used to compress some frequently occurring data into a compressed form while leaving the remaining on the same length. Thus, on the recovering side it produces a probable chances for not to recover the original data with failure probability of Δ. This type of compression is useful in images but does not produce as much reliable compression ratio.

1.2. Lossless Compression

Lossless compression compresses all the files of data on different lengths. Hence, to have an effective compression we make compression of particular files much shorter than their

original size having more probability to occur in the uncompressed files and make lengthy compression for those files have low probability to appear in the original data files.

The idea about the compression which was proposed by the Claude Shannon in its paper (1948) [13], was based on compromising of such elements of an ensemble which are less probable to occur in a series of data stream by introducing some error probability like appearance of ASSCII elements {!,}, {, %, [,]} is less probable in a text file (English Text) which can reduce the size of the ensemble to compress only the less probable data elements. Suppose we have an ensemble A_x with probabilities P_x as, A_x = {a,b,c.d,e,f,g,h}

And, P_x = {3/10, 1/5, 1/5, 1/5, 1/30, 1/30, 1/30}

By leaving 4 less probable elements with their combined probability Δ=2/15, we can reduce code for our ensemble from 3 bits to 2 bits with another ensemble A_Δ which is a subset of A_x i.e. A_Δ= {a, b, c, d} with error probability Δ= 2/15 i.e. $P(x \notin A_\Delta) \leq \Delta$ as shown in table below:

Δ=0		Δ=2/15	
Elements	Codes	Elements	Codes
a	000	a	00
b	001	b	01
c	010	c	10
d	011	d	11
e	100	e	--
f	101	f	--
g	110	g	--
h	111	h	--

Table 1. Lossless Compression

Such data compression approach is also called Entropy Encoding but it introduces ambiguities when each and every codeword is equally likely or having the same probability to occur. Obviously no compromise can be done on any element of the data files that was a problem in Shannon Fano Code (1948) [11, 12] which was compensated by other data compressing coding algorithms like Huffman Code [1], Arithmetic Code [2], Run Length Encoder [8], and Lempel–Ziv–Welch (LZW) [3, 4, 14, 15] which are going to be further discussed in the way proceeding in this paper.

2. COMPARATIVE ANALYSIS OF DIFFERENT DATA COMPRESSING ALGORITHMS

In the sub-sections of this section we will examine different data compressing algorithms and compare them regarding their advantages, disadvantages and applications.

2.1. Shannon Fano Code

Shannon and Fano Algorithm [11, 12] was the first approach in data compression which provided the base for Huffman to design his algorithm with the amendments in Shannon and Fano algorithm. Shannon Fano algorithm was based on:

- Top to bottom approach.
- Elements are assigned with the codeword based on their frequency or probability.
- A repeated procedure is followed to divide each branch into a possible of two states of 0 and 1 until all the elements are assigned with the unique codeword.

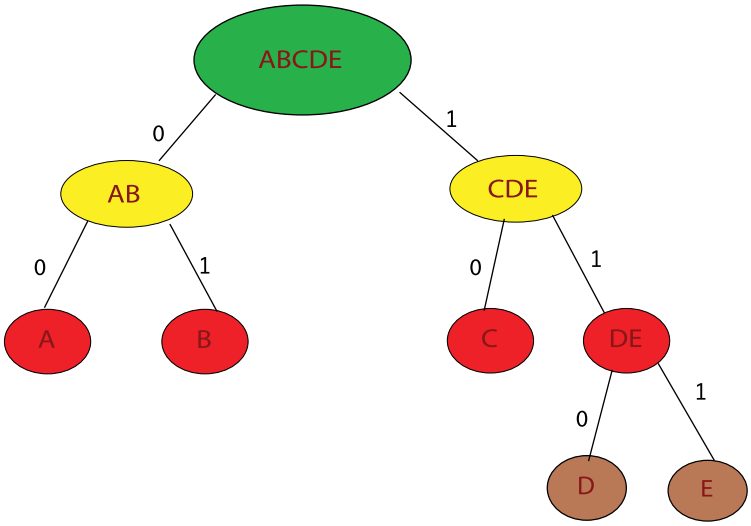


Figure 1. Shannon Fano Algorithm

Elements	Frequency	Code	Code Length
A	10	00	2
B	15	01	2
C	8	10	2
D	7	110	3
E	6	111	3

Table 2. Shannon Fano Coding

Total Elements: 46

Shannon-Fano Code: 105 bits required

Linear code with 3 bits per element: 138 bits required

Shannon-Fano code can be coded with $log_2(105)$ = 6.71 ~ 6bits with some error probability Δ.

Linear code can be coded with $log_2(138)$ =7.10 ~ 7bits.

2.2. Huffman Coding Algorithm

Huffman designed a coding scheme which was based on Shannon-Fano algorithm [11, 12] but it got so much recognition in the field of information theory because such codes were prefix-free codes [7] (each preceding codeword is not a prefix of the other codeword) and assign shortest possible code to the elements having high frequency to appear in the complete data file.

2.2.1. Adaptive Huffman Coding

Adaptive Huffman coding [10] is also known as Dynamic Huffman coding as it is a real time process in which codes are build up when the data is being transmitted, having no any prior knowledge of source distribution. It is beneficial for real time encoding but very sensitive for transmission errors as a single bit error can destroy the whole data transmission.

2.2.2. Application of Huffman Coding Algorithm

- Use as a back end support for different data compressing techniques as PKZIP's algorithm, JPEG (Joint Photographic Experts Group), and MP3.

- Due to its simple and fast implementation, Huffman coding [1] is very preferential in assigning shortest possible codes.

2.2.3. Advantages

- Compact and optimally coded data.

- Fast and easy to use.

- A lossless data compression technique.

2.2.4. Disadvantages

- Files having a repeated pattern of text or image colors can be more effectively compressed by Run Length Encoding [8] than Huffman coding [1].

- System statistic should be known to have an effective Huffman coding.

- An effective Huffman coding [1] is based on two processes. In one a statistical model is build up and in second an image of the encoded data is formed from its statistical model. In view to this approach, it is a slow process as compare to others like Arithmetic coding [2].

- All the codes have different code lengths in Huffman algorithm [1] that makes it difficult for the decoder to realize the end of the bits and the only process left is to follow the up-side down tree to reach the end. If additional bits (corrupted data) will be introduced into data stream so the decoder will be lost and only garbage data will display/print.

- Possible solution for above problem is to send a Huffman table along with the data so that the decoder can decode it properly but this will introduce overheads.

- Huffman coding [1] introduces overhead of 1 bit per code which makes it impossible to compress a data of 1 bit.

2.3. Arithmetic Coding Algorithm

A more generalized view of Huffman coding [1] is the Arithmetic coding [2] which can more easily assign codes to a binary string. Instead of breaking the data into different constituent parts (symbols) representing each with a code as other entropy encoding like Huffman algorithm does, arithmetic coding encodes the whole data with a single number.

2.3.1. Explanation

Encoding of data L, M, N by block coding as L=00, M=01, N=10, will leave the 11 as unused.

- If data LMMNLM is represented in decimal form in base 10, it will be as⟦(0.011201)⟧_10.

- Since arithmetic code has only binary representation with precision 0 to 1 so converting it in base 2 will be as ⟦(0.0010110010)⟧_2 (arithmetic coded data).

- If it is encoded by block code, it will be as LMMNLM = 000101100001. Thus, 2 bits are saved by arithmetic coding [2]. It becomes more effective when we have a long stream of data.

- Arithmetic coding [2] works based on the prediction about the data due to some known patterns. The more accurate the guess, the more optimally encoded data will be generated. Thus, it is an adaptive model.

- Each interval between 0 to 1 divides into sub-intervals as a fractional part of the current intervals with their probabilities associated with the symbols will be in the current context and sub-divisions of the current divisions keep on until a termination statement id is found in data.

2.3.2. Example

Suppose our data have four words with their probabilities or fractional part from 0 to 1 is as:

- **Equal** with fractional value [0,0.5)

- **Greater** with fractional value [0.5,0.7)

- **Less** with fractional value [0.7,0.8)

- **Termination** with fractional value [0.8,1)

And the data is 0.398 to be encoded. Firstly the data 0.398 fall in range 0 to 0.5, so it will be encoded as **Equal** and now further sub-division will be as,

- **Equal** will be 50% of the range [0,0.5) i.e. [0,0.25)
- **Greater** will be 20% of the range [0,0.5) i.e. [0.25,0.35)
- **Less** will be 10% of the range [0,0.5) i.e. [0.35,0.4)
- **Termination** will be 20% of the range [0,0.5) i.e. [0.4,0.5)

In second step the data 0.398 fall in the range 0.35 to0.4, so it will be encoded as Less and further sub-division will be as,

- Equal will be 50% of the range [0.35,0.4) i.e. [0.35,0.375)
- Greater will be 20% of the range [0.35,0.4) i.e. [0.375,0.385)
- Less will be 10% of the range [0.35,0.4) i.e. [0.385,0.30)
- Termination will be 20% of the range [0.35,0.4) i.e. [0.39,0.4)

This time the data 0.398 fall in the range [0.39, 0.4) which will be encoded as **Termination** that means all the data has been encoded.

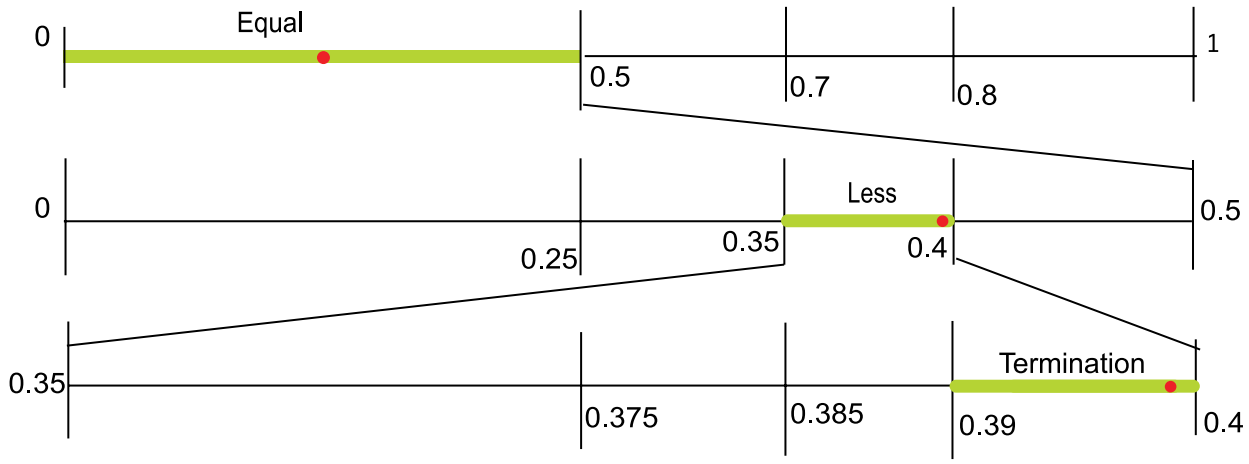


Figure 2. Arithmetic Coding

Note: Due to ease of understanding, I have taken all the data and calculation in base 10 in the range 0 to 1 but for arithmetic coding [2], it should be sub-divided in base 2 (binary) in the range 0 to 1.

2.3.3. Application

- Usually used in the applications where frequently repeating patterns or sequences with known system statistic are used.

2.3.4. Advantages

- A lossless data compressing technique.
- Most suitable in assigning fast and shortest codes for repeating data sequences like English text files with only 2 extra bits added in the compression of entire data sequence.

2.3.5. Disadvantages

- Prior system statistic should need to be known.
- Charges are to be paid to the patent for implementing the algorithm (paid algorithm).

2.4. Run Length Encoding

When frequency of data repetition is very high, so it can be represented by a single item with an integer value representing its number of times repetition in the original sequence of data as explained below with an example:

Data = LLLLLLLLLLMMMLXXXZYX

Above data steam has lots of repetition of alphabets which can be reduced by using Run Length Encoding [8] as,

Data ⟶ (compressed by run length encoding) ⟶ L10M3L1X3Z1Y1X1

Now the original string from 24 characters has been reduced into 17 (20% compression). It can further reduce if we neglect the counter 1 from all those values which are repeating once in the data stream as,

Data ⟶ (compressed by ignoring counter 1) ⟶ L10M3LX3ZYX

Now we have 13 characters out of 24 in original data (46% compression).

2.4.1. Application

Its application is in the file formats like TIFF (Tagged Image File Format), BMP (Bitmap Image) and PCX (Picture Exchange).

2.4.2. Advantages

- A lossless data compression technique.
- Very easy, simple and fast to implement.

2.4.3. Disadvantages

- Compression ratio depends on the type of the data file and remains low for most of the cases when comparing to other algorithms.

There are two effective transformations which are used in combination with the coding algorithms like Huffman [1] and Arithmetic coding [2] to make the data compressing algorithm techniques much more effective in their applications. Besides this, they are fast and easy to implement. These are:

1. Move-to-Front Transform
2. Burrows and Wheeler Transform (BWT)

2.5. Move-to-Front Transform

Each byte of data is encoded by the index value assigned from all the values of elements (alphabets) in the list. More deep understanding can be developed by throwing light on an example below:

2.5.1. Example

In this example we are assigning index values to data (CABCBBB) from a to z alphabets from all the possible data values in the list. On assigning an index value to each element, it is move to the front of the list and the process goes on for the next coming values.

2.5.2. Encoding

Data	Sequence	List	Next Action Required
CABCBBB	2	ABCDEFGHIJKLMNOPQRS TUVWXYZ	In the next step C will move to the front in the list.
CABCBBB	2,1	CABDEFGHIJKLMNOPQRS TUVWXYZ	In the next step A will move to the front in the list.
CABCBBB	2,1,2	ACBDEFGHIJKLMNOPQRS TUVWXYZ	In the next step B will move to the front in the list.
CABCBBB	2,1,2,2	BACDEFGHIJKLMNOPQRS TUVWXYZ	In the next step C will move to the front in the list.
CABCBBB	2,1,2,2,1	CBADEFGHIJKLMNOPQRS TUVWXYZ	In the next step B will move to the front in the list.
CABCBBB	2,1,2,2,1,0	BCADEFGHIJKLMNOPQRS TUVWXYZ	In the next step B will move to the front in the list.
CABCBBB	2,1,2,2,1,0,0	BCADEFGHIJKLMNOPQRS TUVWXYZ	
Final	2,1,2,2,1,0,0	BCADEFGHIJKLMNOPQRS TUVWXYZ	

Table 3. RLE Encoding

2.5.3. Decoding

Data	Sequence	List	Next Action Required
2,1,2,2,1,0,0	C	ABCDEFGHIJKLMNOPQRSTU VWXYZ	In the next step C will move to the front in the list.
2,1,2,2,1,0,0	CA	CABDEFGHIJKLMNOPQRSTU VWXYZ	In the next step A will move to the front in the list.
2,1,2,2,1,0,0	CAB	ACBDEFGHIJKLMNOPQRSTU VWXYZ	In the next step B will move to the front in the list.
2,1,2,2,1,0,0	CABC	BACDEFGHIJKLMNOPQRSTU VWXYZ	In the next step C will move to the front in the list.
2,1,2,2,1,0,0	CABCB	CBADEFGHIJKLMNOPQRSTU VWXYZ	In the next step B will move to the front in the list.
2,1,2,2,1,0,0	CABCBB	BCADEFGHIJKLMNOPQRSTU VWXYZ	In the next step B will move to the front in the list.
2,1,2,2,1,0,0	CABCBBB	BCADEFGHIJKLMNOPQRSTU VWXYZ	
Final	CABCBBB	BCADEFGHIJKLMNOPQRSTU VWXYZ	

Table 4. RLE Decoding

(CABCBBB) is the data retrieved after decoding through move-to-front transform [9] which is same as the original data. Since final list is same for both the encoded and decoded data which shows that no any bit was flipped during the transmission. That's why we have retrieved the original data safe and soundly.

2.6. The Burrows-Wheeler Transform

Burrows-Wheeler Transform [5] is a way to sort data into matrix of N*N which is a lossless and reversible transform. The best part of BWT is that it can convert the data into a form that is more feasible to deal by other compressing techniques more efficiently and inverse BWT [6] can retrieve the data in its original form. BWT transformed data is more compressible which was invented by Michael Burrows and David Wheeler in 1994. Basically BWT was a solution for drawbacks of run length encoding [8]. In run length encoding [8], we transformed the more repeated text with an index to show its repetition in original data but the data usually does not adopt such a form in which the repetition in text occurs so often. So BWT transforms the data in the form which accumulates the repeated alphabets on a single place in the last column of the transformed matrix so that they could be more flexibly represented by run length encoding [8] with more efficient compressed form. By this way, BWT becomes more efficient (more data compression) with the combination of locally-adaptive algorithms such as run length encoding [8] or move-to-front coding [9] in combination with Huffman [1] or arithmetic coding [2].

2.6.1. Steps to Implement BWT

Step 1

Append a $ as the last character in the text data as a sentinel.

Step 2

Form a N*N matrix with the given data by cyclically shifting the last character as a first one in the next row of the matrix.

Step 3

Arrange the data in ascending order row-wise.

Step 4

The transformed data is the last column together with the primary index of the row where the original data lies.

2.6.2. Example

Let suppose the given text data is,

T = mississippi

Step 1:

T$ = mississippi$

Step 2:

m	i	s	s	i	s	s	i	p	p	i	$
i	s	s	i	s	s	i	p	p	i	$	m
s	s	i	s	s	i	p	p	i	$	m	i
s	i	s	s	i	p	p	i	$	m	i	s
i	s	s	i	p	p	i	$	m	i	s	s
s	s	i	p	p	i	$	m	i	s	s	i
s	i	p	p	i	$	m	i	s	s	i	s
i	p	p	i	$	m	i	s	s	i	s	s
p	p	i	$	m	i	s	s	i	s	s	i
p	i	$	m	i	s	s	i	s	s	i	p
p	$	m	i	s	s	i	s	s	i	p	p
$	m	i	s	s	i	s	s	i	p	p	i

Figure 3. BWT Transform Data Initialization

Step 3:

Figure 4. BWT Transform Data Sorting

Note: Notice that the transformed data (last column of transformed matrix) has now 2s's together twice and 3i's together once which make the compression through run length coding [8] more efficient (more compressible).

Step 4:

- Primary Index = 4
- L= ssmp$pissiii

2.6. The Burrows-Wheeler Transform

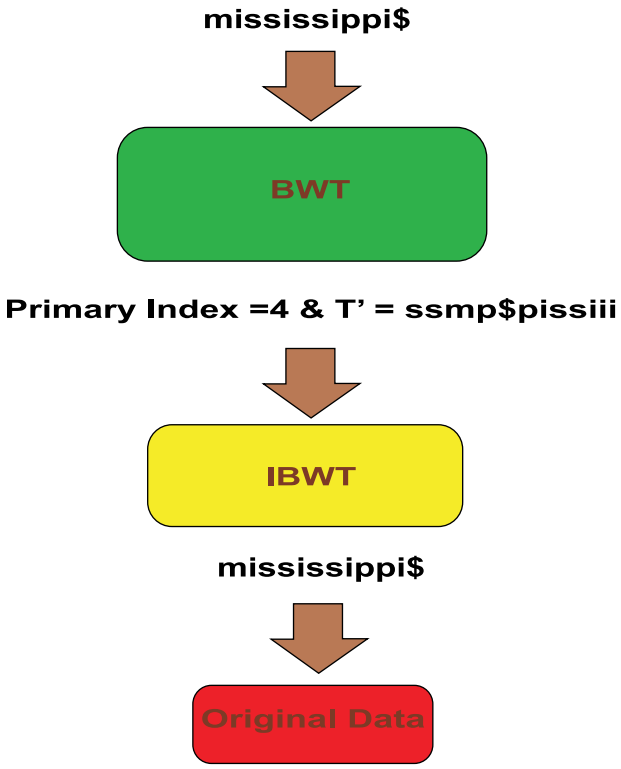


Figure 5. IBWT Transform

In order to retrieve the original data by applying Inverse Burrows-Wheeler Transform [6], we use LF mapping.

2.6.4. LF Mapping

Step 1

Pick up the first column in transformed matrix as F.

Step 2

Pick up the last column in transformed matrix as L.

Step 3

Based on F and data in primary index that is 4 in our case, assign index numbers to data in last column.

Step 4

Generate the original data by applying LF mapping on the element positions presented in the last column of transformed matrix from their index numbers which were assigned in the previous step, starting first from primary index and then for the number obtained from previous LF mapping of primary index and so on.

2.6.5. Example

Step 1

F = iiimppssss$

Step 2

L = ssmp$pissiii

Step 3

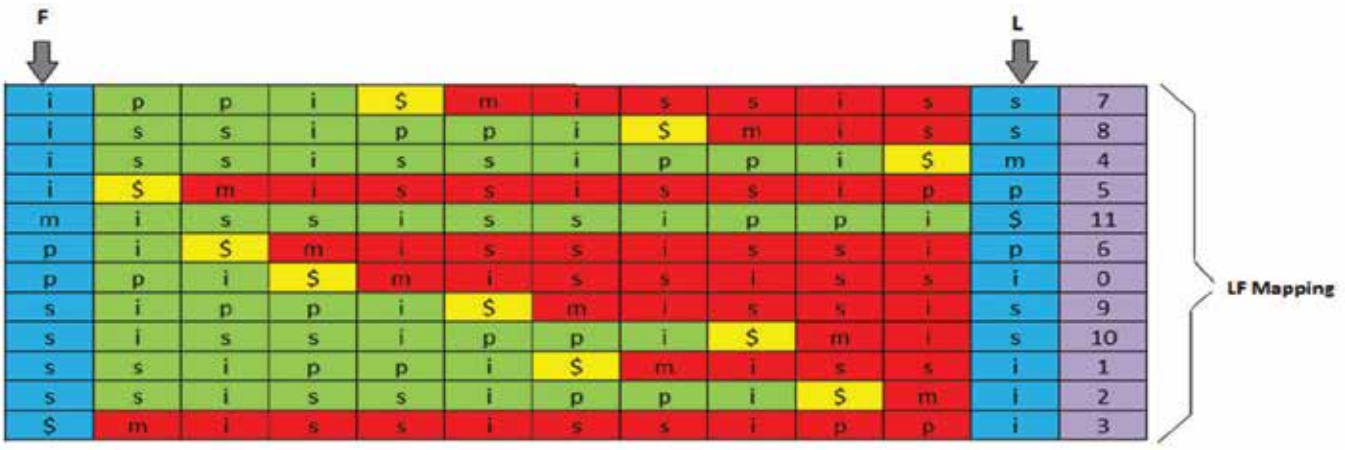


Figure 6. LF Mapping

Step 4

- Primary Index s = 4
- s = LF[4] =11 T = [----------$]
- s = LF[11] = 3 T = [----------i$]
- s = LF[3] = 5 T = [----------pi$]
- s = LF[5] = 6 T = [---------ppi$]
- s = LF[6] = 0 T = [--------ippi$]
- s = LF[0] = 7 T = [-------sippi$]
- s = LF[7] = 9 T = [------ssippi$]
- s = LF[9] = 1 T = [-----issippi$]
- s = LF[1] = 8 T = [----sissippi$]
- s = LF[8] = 10 T = [---ssissippi$]
- s = LF[10] = 2 T = [--ississippi$]
- s = LF[5] = 4 T = [mississippi$]

2.6.6. Application

- Applicable with all those data compressing techniques which require more data repetition for their encoding.

2.6.7. Advantages

- Transformed text provides more probability for getting the identical characters in more vicinity.
- The transformed data can be retrieved properly from IBWT.

2.6.8. Disadvantages

- BWT is a super-linear time complex as the data needs time to get its transformed form.
- Not hardware friendly software because of its super-linear time complexity.
- It works only for a block of complete text data at a time and cannot be implemented on the whole text data directly.

2.7. Lempel–Ziv Welch (LZW) Compression

Lempel-Ziv Welch [3, 4, 14, 15] is not a static compression technique that does not require prior knowledge before encoding and decoding of data. It maintains its own dictionary based on data string, that's why it is also called universal coding scheme which was invented by Abraham Lempel, Jacob Ziv and Terry Welch. This is the technique used in PC's to double their hard drive capacity. LZW compression consists of a table of 4096 entities from which 0 to 255 are consumed by the ASCII characters to maintain an initial table and codes from 256 to 4095 are assigned to all those entities which do not repeat in the data sequence.

2.7.1. LZW Encoding Steps

1. Initialize a table by assigning ASCII codes to the given data sequence.

2. Initially P will be empty.

3. C is the next character to introduce in data sequence.

4. In the next step P=P+C of previous step.

5. Maintain dictionary by adding P+C.

6. Output for each P+C will be the ASCII code of the alphabet in C.

7. Repeat all the steps from 1 to 6 until you get the last entity of data sequence.

2.7.2. Example:

Data= ABRACADABRA

P=Empty & C=A

C	P	P+C	Is seen before?	Index	Dictionary	Output
xxxxx	xxxxx	xxxxx	Yes	1	R	xxxxx
xxxxx	xxxxx	xxxxx	Yes	2	D	xxxxx
xxxxx	xxxxx	xxxxx	Yes	3	C	xxxxx
xxxxx	xxxxx	xxxxx	Yes	4	B	xxxxx
A	xxxxx	A	Yes	5	A	xxxxx
B	A	AB	No	6	AB	5
R	B	BR	No	7	BR	4
A	R	RA	No	8	RA	1
C	A	AC	No	9	AC	5
A	C	CA	No	10	CA	3
D	A	AD	No	11	AD	5
A	D	DA	No	12	DA	2
B	A	AB	Yes	xxxxx	xxxxx	xxxxx
R	AB	ABR	No	13	ABR	6
A	R	RA	Yes	xxxxx	xxxxx	xxxxx
End of String	RA	RA	Yes	xxxxx	xxxxx	8

Encoded Data= 5 4 1 5 3 5 2 6 8

Table 5. LZW Encoding

Note: For the ease of understanding in the above example, I have used index numbers instead of the ASCII codes for different characters. But in reality they are represented by their ASCII codes as machines are usually programmed in 8 bit ASCII codes from 0 to 255.

2.7.3. LZW Decoding Steps

1. Initialize a table by assigning ASCII codes to the encoded data sequence.

2. Assign the characters based on their ASCII codes in CW to string.cw.

3. Initially string.pw and PW will be empty and then get the values of string.cw and CW from previous row in the table as the current entities of string.pw and PW respectively.

4. Maintain dictionary by adding P+C.

5. Output will be the ASCII characters in string.cw.

6. Repeat the steps from 1 to 5 until you get the last encoded data code.

2.7.4. Example

Encoded data= 5 4 1 5 3 5 2 6 8

CW	PW	P	C	P+C	string.pw	string.cw	Index	Is seen before?	Dictionary	output
xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	1	Yes	R	xxxxx
xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	2	Yes	D	xxxxx
xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	3	Yes	C	xxxxx
xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	4	Yes	B	xxxxx
5	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	A	5	Yes	A	A
4	5	A	B	AB	A	B	6	No	AB	B
1	4	B	R	BR	B	R	7	No	BR	R
5	1	R	A	RA	R	A	8	No	RA	A
3	5	A	C	AC	A	C	9	No	AC	C
5	3	C	A	CA	C	A	10	No	CA	A
2	5	A	D	AD	A	D	11	No	AD	D
6	2	D	AB	DAB	D	AB	xxxxx	Yes	xxxxx	AB
8	6	AB	RA	ABRA	AB	RA	xxxxx	Yes	xxxxx	RA

Output= Decode data= ABRACADABRA= Original Data

Table 6. LZW Decoding

Note: For the ease of understanding in the above example, I have used index numbers instead of the ASCII codes for different characters. But in reality they are represented by their ASCII codes as machines are usually programmed in 8 bit ASCII codes from 0 to 255.

2.7.5. Application

- Most frequently apply in GIF (Graphic Interchange) and TIFF (Tagged Image File Format).

2.7.6. Advantages

- The most efficient coding algorithm to compress data based on repeating sequences.
- Does not require any prior statistic of the system (data sequence).
- Most fast in execution.
- Lossless data compression.

2.7.7. Disadvantages

- It has its limitation to introduce only 256 to 4096 entities. After that dictionary needs to be re-build again.
- LZW is a paid data compression technique.
- Data table is difficult to maintain.

2.7.8. Application

- Its application is in the file format like GIF (Graphics Interchange Format).

3. CONCLUSION

Efficient video and image compression is now not a big issue with the advancements in data compressing techniques which have been discussed in detail in this paper. Choice of particular compressing techniques is basically based on the requirements of the users. Each has its own particular advantages preferable in particular applications that's why each one is in use but more advancements could be possible by some other advanced techniques still remain dormant to appear in future that can provide more fast, easy and efficient data compression with fulfillment of all possible requirements of the users. Below is a table of comparison creating from all the stuff discussed in detail about different data compressing techniques:

Algorithm	Huffman	Arithmetic	RLE	BWT	LZW
Advantages	A lossless data compression technique.	A lossless data compressing technique.	A lossless data compression technique.	Implement on top of the both lossy and lossless data compressing techniques	Lossless data compression. Does not require any prior statistic of the system (data sequence).
	Fast and easy to use. Compact and optimally coded data.	Most suitable in assigning fast and shortest codes for repeating data sequences like English text files with only 2 extra bits involved in the compression of entire data sequence.	Very easy, simple and fast to implement, requires few hardware elements and software resources.	Transformed text provides more probability for getting the identical characters in more vicinity. Transformed data can be retrieved properly from IBWT.	Most fast in execution. The most efficient coding algorithm to compress data based on repeating sequence.

Disadvantages	System statistic should be known to have an effective Huffman coding. Files having a repeated pattern of text or image colors can be more effectively compressed by Run Length Encoding than Huffman coding. Huffman coding introduces overhead of 1 bit per code which makes it impossible to compress a data of 1 bit.	Prior system statistic should need to be known.	Compression ratio depends on the type of the data file and remains low for most of the cases when comparing to other algorithms.	It works only for a block of complete text data at a time and cannot be implemented on the whole text data directly.	It has its limitation to introduce only 256 to 4096 entities. After that dictionary needs to be re-build again.
	All the codes have different code lengths in Huffman algorithm that makes it difficult for the decoder to realize the end of the bits and the only process is to follow the up-side down tree to reach the end. If additional bits (corrupted data) will be introduced into data stream so the decoder will be lost and only garbage data will display/print. Possible solution to the problem is to send a Huffman table along with the data so that the decoder can decode it properly but this will introduce overheads.	Charges are to be paid to the patent for implementing the algorithm (paid algorithm).	This type of encoding is only suitable to compress the data with recurring or identical symbols.	BWT is a super-linear time complex as the data wants time to get its transformed form. Not hardware friendly software because of its super-linear time complexity.	LZW is a paid data compression technique. Data table is difficult to maintain.

Applications	Using as a back end support to different data compressing techniques as PKZIP's algorithm, JPEG (Joint Photographic Experts Group), and MP3. Due to its simple and fast application, Huffman coding is very preferential in assigning shortest codes.	Usually used in the applications where frequently repeating patterns or sequences with known system statistic are used.	Its application is in the file formats like TIFF (Tagged Image File Format), BMP (Bitmap Image) and PCX (Picture Exchange).	Best suitable to apply with all those data compressing techniques which requires more data repetition for their encoding.	Its application is in the file format like GIF (Graphics Interchange Format).

Table 7. Comparison of Different Coding Algorithms

REFERENCES

[1] A Concise Introduction to Data Compression, Saloman, David, 2008, XIV, 314 p. 89 illus, with online files/ update, softcover, ISBN: 978-1-84800-071-1

[2] Arithmetic Coding for Data Compression, Ian H. Witten, Rodford M. Neal, & John G. Cleary, Communication of the ACM, June 1982, Volume 30, Number 6

[3] The Lemper Ziv Algorithm, Christina Zeeh, Seminar "Famous Algorithms", January 16, 2003

[4] https://www.cs.duke.edu/csed/curious/compression/lzw.html

[5] A Simple Analysis of Burrows-Wheeler-Based Compression, Haim Kaplan, Shir Landau, Elad Verbin, Elsevier (Theoretical Computer Science) 308 (2007) 220 235

[6] Functional Pearls, Inverting the Burrows-Wheeler Transform, Richard Bird & Shin-Cheng MU, under consideration for publication in J. Fundamental Programming

[7] https://en.wikipedia.org/wiki/Prefix-Code

[8] https://en.wikipedia.org/wiki/Run-length-encoding

[9] https://en.wikipedia.org/wiki/Move-to-front_transform

[10] https://www.icg.isy.liu.se/en/couses/tsbk08/lect4.pdf

[11] https://en.wikipedia.org/wiki/Shannon-Fano-Coding

[12] http://www.binaryessence.com/dct/en000046.htm

[13] A Mathematical Theory of Communication, C.E Shannon, The Bell System Technical Journal Vol. 27, pp. 379-423, 623-656, July, October, 1948

[14] https://courses.cs.washington.edu/courses/cse326/05au/lectures/lecture19.pdf

[15] Compression Algorithms: Huffman & Lempel-Ziv-Welch (LZW), MIT 6.02 Draft Lecture Notes, February 13, 2012

ANALYSIS OF A QUEUEING MODEL FOR PAYLOAD DATA PROCESSING IN A UAV REMOTE SENSING ENVIRONMENT

Anas Bin Iftikhar, Sameer Qazi, Fahd Khan, Arsalan Jawed
Karachi Institute of Economics and Technology
Email: {anas.iftikhar, sameer.qazi, fahd.khan, arsalan.jawed}@kiet.edu.pk

Received: 26-November-2022 / Accepted: 12-December-2022
Karachi Institute of Economics and Technology || Technology Forces Journal, Volume 4, Issue 2, 2022

ABSTRACT

UAV platforms are showing renewed interest in the field of remote sensing, leveraged by their high maneuverability and reduced preparation time. However, the on-board data processing is still a challenge. Typical UAV missions record data and are then processed after the flight, increasing the data delivery from days to weeks. This is prohibitive in situations where real-time data analysis is required for fast decision making in remote sensing and SAR missions. With the advancement of embedded computing systems and multi-core architectures, remote sensing can benefit from online data processing and faster result delivery. There, however, remains a challenge for system designers to use hardware resources efficiently. This study aims to explore an existing data processing software architecture using queuing model approach and proposes a better architecture with reduced latency and increased throughput.

Index Terms: UAV, Remote Sensing, Queue, FIFO, Determin- istic, Markov, Stochastic, Search and Rescue

1. INTRODUCTION

The applications for unmanned aerial vehicles (UAVs) are growing rapidly in civil and commercial sector, owing to their ease of deployment, low maintenance cost, increases mobility and precision maneuverability [1] [2]. With the recent development of improved sensing technologies along with precision guidance and navigation algorithms, UAV's are now finding excellent opportunities in the civil and commercial market [3]. Presently, UAV's are being used extensively in a remote sensing setting with applications such as precision agriculture, multi-spectral imaging, traffic monitoring, public security, industrial inspection and media coverage [4].

In a typical UAV remote sensing scenario, the raw sensor data is often processed to get the final results. In geo-spatial sensing applications, the imagery acquired through the on- board optical sensors and imagers are required to be stitched together to get a single ortho-rectified image [5]. In surveil- lance and security applications, moving target and tracking is required through computer vision and image processing [6]. In industrial inspection applications the detection of surface defects and structural failures is required which is often achieved by employing computationally intensive algorithms [7]. The search and rescue operations using UAV platforms require strong situational and context

awareness for successful missions [8]. The data acquired from the UAV platforms often require accurate and precise geo-referencing, which also demands compute intensive algorithms such as real-time kinematics and post-processing kinematics [9].

Majority of such applications in the existing deployment ready UAV systems require post-processing before any final result can be delivered. Several solutions exist in the industrial applications and often utilize cloud-processing or dedicated processing server equipped with hardware acceleration (GPU, FPGA) [10]. While such solutions offer greater precision and data quality, the turn around time for final result delivery can often increase exponentially. Many applications such as search and rescue, precision agriculture and industrial inspection require processing to be carried out on-board which introduces myriad of challenges [11].

UAVs equipped with multiple sensors and payloads require a sensor-fusion to create a situational awareness picture, es- pecially for search and rescue missions. Utilizing dedicated processing hardware for individual sensor stream becomes prohibitive in terms of UAV SWaP (size, area and power) [12]. Using parallel processing hardware such as GPU and FPGA require an extra power capital from UAV power distribution which affects the overall power consumption, thereby affecting UAV's flight time. For vehicles designed for long endurance search and rescue mission, it is imperative that the data processing hardware must consume less power while still producing desired outputs.

Multi-core processors on small embedded systems are now ubiquitous. Such systems also incorporate floating-point processing nodes as well as on-chip hardware accelerators for matrix operations. Despite such option on the hardware level, a corresponding software architecture is also required to handle multiple processing tasks and shared resources. Typical programming languages are inherently sequential in nature and thus becomes challenging for system designers to realize task parallelism.

In this study, an existing processing scheme is analyzed using queuing model. After this, a novel software architecture is proposed which uses multiple queues and task parallelism to speed up the execution and improve performance.

2. OVERALL SYSTEM DESCRIPTION

The system under study is a search and rescue UAV platform in a VTOL fixed-wing configuration equipped with visible and infrared camera payloads along with multi GNSS system for accurate position information. The functional requirement of the system is to transmit a single situational awareness video with each video frame consisting of:

A Thermal hot-spots from acquired IR images.

B Moving object detection with bounding boxes, acquired from visible camera.

C An overlay algorithm to fuse both images acquired into a single image.

The entire situational-awareness system comprises of sub- systems as follows. Two camera payloads with a Gigabit- Ethernet interface and video encoder, a central data-processing node based on quad-core Raspberry Pi4 with Gigabit-Ethernet port, 4GB DDR4 RAM, H264 video encoder/decoder and a 2.8GHz radio data-link based on OFDM scheme. The raw H264 video stream is acquired from the ground station played on the media players as an RSTP protocol.

A. Algorithm Description

The embedded data processing system consists of three pro- cessing tasks. The first task is a hot-spot detection algorithm which acquires video stream from thermal camera payload. The video stream is a monochromatic, gray-scale where each pixel of the image is already referenced as temperature inten- sity. The hot-spot detection algorithm searches for maximum pixel values as a subset of entire 2D image. This is achieved by sweeping the entire image using a window size of 2-by-2 pixels and a step size of 1. This new

subset is then reshaped and corresponding pixel value locations in the original image are found. These locations along with the pixel value are then used to form a bounding box around the hot-spot.

The second task detects moving object from visible day camera using optical flow algorithm based on the Horn- Schunck method [13]. The optical flow algorithm used here is an iterative algorithm which finds the direction of moving

pixel as an optimization problem. This algorithm is robust to distortions from lens and camera movement, incurred from the UAV movement.

The third task acquires input from both hot-spot and motion detection processes and renders a single frame, consisting of bounding boxes of moving objects as well as hot-spot objects. A system-level block diagram is shown in the figure 1.

B. Algorithm Execution Time Analysis

The execution time of each task is empirically calculated by running the entire application and processing 500 frames. The mean execution time μ, standard deviation σ along with coef-ficient of variation cv are calculated. The variable cv demon-strates the relative dispersion of execution time around the mean. The mean execution time and coefficient of variability are shown in the Table I. It is worth noting that the coefficient of variability is very close to zero, demonstrating that the processing tasks are following a deterministic behaviour (near real-time).

The existing architecture of processing chain is imple- mented as a sequence of operations, where each node is processed sequentially. The system is realized as a queuing model, where incoming images are treated as incoming clients and tasks are considered as services. The activity diagram of the existing processing chain is shown in the Figure 1. The values of interest are average time in the system (W),

TABLE I. MEAN AND COEFFICIENT OF VARIATION OF THE EXECUTION TIME

Processing Task	Mean Execution Time (in seconds) (τ)	Coefficient of Variation (cv)
Optical Flow	0.246	0.039
Ho$^{\tau}$-Spd$^{\tau}$	0.186	0.038
Overlay	0.09	0.04

average waiting time (Wq), average elements in the system (L), average elements in the queue (Lq) and server utilization (ρ).

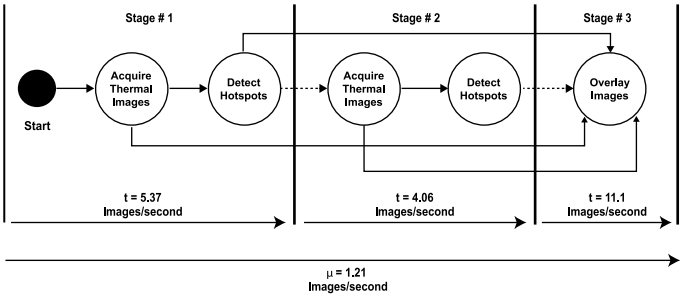


Fig. 1. System Block Diagram

C. Server Utilization and System Performance Analysis

For each stage, a D/D/1 queue model is considered with average arrival rate, $E(A) = \frac{1}{\lambda}$ and departure rate $E(S) = \frac{1}{\mu}$.

The average elements in the system (L), and average waiting time (W) are given as Eq 1 and Eq 2 respectively. It is worth noting that due to deterministic arrival and departure rates, the average waiting time and average elements in the queue are always zero. By varying the λ and μ, the server utilization can assume any value between 0 and 1.

$$L = \rho = \frac{\lambda}{\mu} \quad (1)$$

$$w = E(S) = \frac{1}{\mu} \quad (2)$$

$$L_q = W_q = 0 \quad (3)$$

The acquisition nodes (thermal and visible camera) acquire single image per iteration. It is also worth noting that the latency of entire chain is linearly dependent on the most compute intensive

processing task. Due to this, the arrival rate of individual camera nodes also scale to this factor. The existing scheme has following characteristics:

a Stage 1 is the start of loop with thermal camera acquisi- tion, with an arrival rate of 10 images/ second. The hot- spot service has departure rate of 5.37 images/second. This departure becomes arrival to stage 3.

b Stage 2 starts with visible camera acquisition with 10 images/second arrival rate. The optical flow algorithm has departure rate of 4 images/ second. The departure here also becomes arrivals in the stage 3.

c Stage 3 takes input from both the previous stages. The arrival rate depends on the departure rate of stage 2, with 4 images/second. Since the output from stage 1 is also being fetched in this stage, the output from stage 1 are also forced to be arriving 4 images/second instead of 5 images/ second.

The cumulative execution time of the entire chain is the sum of individual processing nodes, with the latency of 0.52seconds. This latency scales linearly with the service time of individual nodes and the most compute intensive node (optical-flow). By applying D/D/1 model, we arrive at the values of interest depicted in Table II.

TABLE II. SERVER UTILIZATION AND SYSTEM PERFORMANCE

Processing Task	L	L_q	W	W_q	ρ
Optical Flow	2.413	0.00	0.246	0.00	2.413
Hot-Spot	1.87	0.00	0.186	0.00	1.87
Overlay	0.89	0.00	0.09	0.00	0.89
Total	5.173	0.00	0.522	0.00	5.173

3. PROPOSED SOFTWARE ARCHITECTURE FOR QUEUING MODEL

A. System Description

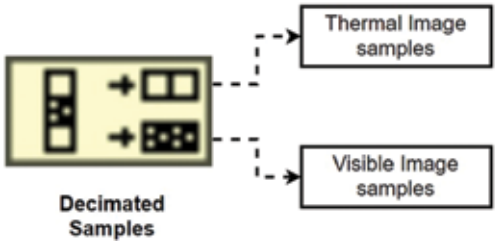
The objective of this study is to reduce the overall latency of the system by introducing task parallelism. This is done by processing individual tasks in parallel, leveraging the multi-core architecture of the embedded processor. The proposed queue network is implemented by using real-time First-in, First-Out buffers to share data between parallel loops. The use of RT FIFOs enable deterministic inter-arrival rate. This makes the proposed network a D/D/S model, where S are the number of service nodes. The proposed network has following characteristics:

a A high rate acquisition loop samples video frames from both payloads at a sampling rate of 200 frames-per- second (200 Hz). These data samples are then interleaved to form a single sensor queue of size 100 elements (Figure 2(b)).

b Individual processing loops (or tasks) decimates incoming sensor queue and forms a local queue. The decimator divides the sensor FIFO, placing output into individual queues successively. The arrival rate of each decimated local queue is $\lambda_t = \lambda_v$ = 10 images per second (Figure 2(a)).

c Each individual loop performs a service at a fixed service time and pushes the output into another local queue. The output of the optical-flow and hot-spot processes are marked as T′ and V′ respectively.

d The output local queues are interleaved again in to a global queue which is then consumed by the overlay process. This process takes successive T′ and V′ elements and forms a single overlaid image.



(b) interleave

Fig. 2. Local Queuing Model

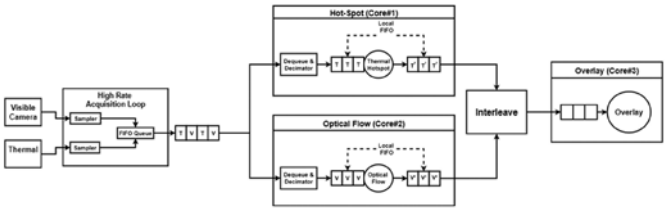


Fig. 3. Proposed Queue Network Model

B. System Performance and Results

The proposed network is characterized by determining av- erage time in the system (W), average waiting time(Wq), average clients in the system (L), average elements in the queue(Lq) and server utilization (ρ). It is an established fact that for arrival rates less than service rates ($\lambda < \mu$), there in effectively no waiting in the D/D/1 queue [14]. By varying the decimation rate, the arrival rate to each processing node can be controlled. The processing time is deterministic but waits for the incoming queue to valid (the queue element holds a valid data). We can, therefore, conclude that the arrival rate should be less than the departure rate to achieve Wq and Lq = 0.

The throughput of the network can be determined by observing the activity diagram shown in the Figure 4. The first stage executes both the optical-flow and hot-spot algorithms in parallel. Hence, the departure rate of this stage is upper limited to the data processing time of most compute intensive task (optical-flow in our case). It is evident that the total throughput increases to 15.17 images/second as opposed to previous architecture where throughput was 1.21 images/second. This staggering increase is due to parallel implementation of both algorithms.

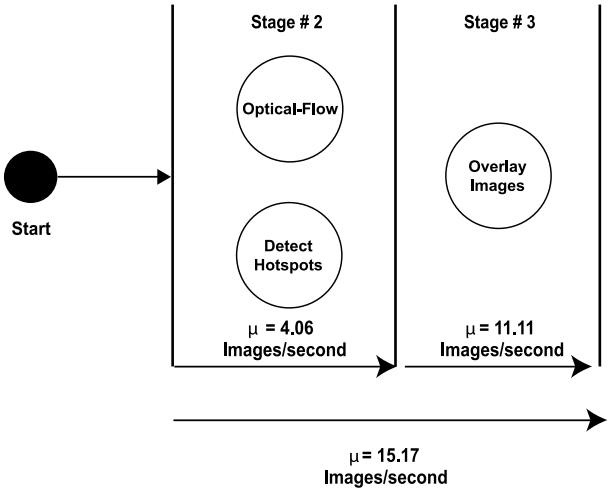


Fig. 4. Throughput of Proposed Network

By keeping the arrival rates less than or equal to the de- parture rate, the number of samples (images) in the queue are equal to zero. In the proposed system, the arrival rate for each local queue is 10 images/second and the departure rate is 4.06 images/second. The server utilization for stage2 is ρ = 2.46.

For optimum performance, the server utilization should not exceed the value of 1. To achieve this, the decimation factor is reduced to match the departure rate of processing nodes i.e. the arrival rates of local queues are set to 4 images/second. A typical server utilization is shown in the Figure 5.

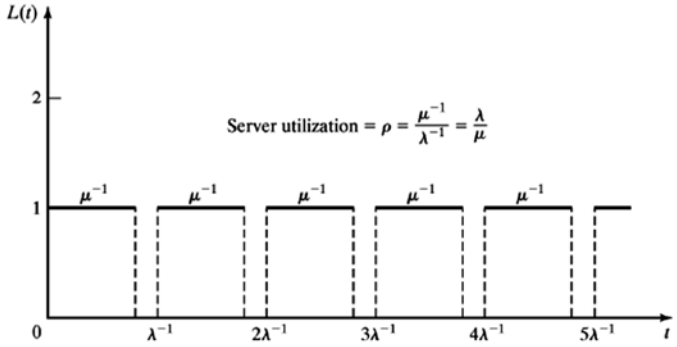


Fig. 5. Server Utilization Graph

By increasing the arrival rate to values greater than the de-parture (or service rate) of processing nodes, a sharp increase in the number of elements in queue Lq is observed. This can be intuitively explained by the graph in Figure 6. Increasing $\lambda > \mu$, the FIFO buffers fills faster than it is being emptied. If the FIFO buffer length is kept constant, the new samples are flushed away, causing data loss. If the FIFO length is kept indefinite, the queue memory reaches its maximum value and causes memory crash. The FIFO schedulers in real-time operating systems are designed to handle this issue by flushing samples or not allowing further samples to enter the FIFO. In either case, the data-loss is unavoidable.

4. FUTURE WORK

This study explored how queuing model approach can be used to design an embedded data processing software. In many real-world UAV applications, the incoming data samples from sensors and payloads often get corrupted due to various

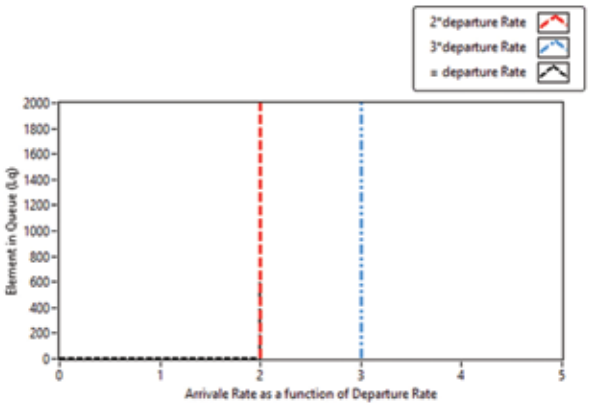


Fig. 6. Queue Length with varying Departure Rate

reasons. The invalid data cannot be processed and thus the software requires error handling. This can be modelled by in- corporating priorities in the queuing network. such processing is critical in applications which are time and safety critical and efficient use to data processing is required to minimize UAV's power consumption.

5. CONCLUSION

In this study, a UAV payload data processing is studied and modelled. First, an existing data processing scheme is analyzed by realizing it as a queue-model. The throughput and latency are analyzed and then a new processing architecture is proposed, leveraging real-time FIFO buffers processor paral- lelism. The network is analyzed using a D/D/S deterministic queuing model, increasing throughput by a factor of 15.

The D/D/S queuing systems are mathematically trivial to model due to their constant inter-arrival and departure rates. However, their implementation in cyber-physical systems is a challenge, requiring efficient schedulers and software architec- tures. This study aimed to explore this challenge for a UAV data processing task and acquired some promising results.

REFERENCES

[1] G. Zhou, V. Ambrosia, A. J. Gasiewski, and G. Bland, "Foreword to the special issue on unmanned airborne vehicle (uav) sensing systems for earth observations," IEEE Transactions on Geoscience and Remote Sensing, vol. 47, no. 3, pp. 687–689, 2009.

[2] H. Shakhatreh, A. H. Sawalmeh, A. Al-Fuqaha, Z. Dou, E. Almaita,

I. Khalil, N. S. Othman, A. Khreishah, and M. Guizani, "Unmanned aerial vehicles (uavs): A survey on civil applications and key research challenges," IEEE Access, vol. 7, pp. 48 572–48 634, 2019.

[3] G. Pajares, "Overview and current status of remote sensing applications based on unmanned aerial vehicles (uavs)," Photogrammetric Engineer- ing & Remote Sensing, vol. 81, no. 4, pp. 281–330, 2015.

[4] C. Zhang and J. M. Kovacs, "The application of small unmanned aerial systems for precision agriculture: a review," Precision agriculture, vol. 13, no. 6, pp. 693–712, 2012.

[5] F. Nex and F. Remondino, "Uav for 3d mapping applications: a review,"

Applied geomatics, vol. 6, no. 1, pp. 1–15, 2014.

[6] L.-Y. Lo, C. H. Yiu, Y. Tang, A.-S. Yang, B. Li, and C.-Y. Wen,

"Dynamic object tracking on autonomous uav system for surveillance applications," Sensors, vol. 21, no. 23, p. 7888, 2021.

[7] Z. A. Siddiqui and U. Park, "A drone based transmission line compo- nents inspection system with deep learning technique," Energies, vol. 13, no. 13, p. 3348, 2020.

[8] R. Geraldes, A. Goncalves, T. Lai, M. Villerabel, W. Deng, A. Salta,

K. Nakayama, Y. Matsuo, and H. Prendinger, "Uav-based situational awareness system using deep learning," IEEE Access, vol. 7, pp. 122 583–122 594, 2019.

[9] H. Zhang, E. Aldana-Jague, F. Clapuyt, F. Wilken, V. Vanacker, and

K. Van Oost, "Evaluating the potential of post-processing kinematic (ppk) georeferencing for uav-based structure-from-motion (sfm) pho- togrammetry and surface change detection," Earth Surface Dynamics, vol. 7, no. 3, pp. 807–827, 2019.

[10] Z. Xu and D. Zhu, "High-resolution miniature uav sar imaging based on gpu architecture," in Journal of Physics: Conference Series, vol. 1074, no. 1. IOP Publishing, 2018, p. 012122.

[11] D. Hulens, J. Verbeke, and T. Goedem e´, "Choosing the best embedded processing platform for on-board uav image processing," in International joint conference on computer vision, imaging and computer graphics. Springer, 2015, pp. 455–472.

[12] A. Montanari, F. Kringberg, A. Valentini, C. Mascolo, and A. Prorok, "Surveying areas in developing regions through context aware drone mobility," in Proceedings of the 4th ACM Workshop on Micro Aerial Vehicle Networks, Systems, and Applications, 2018, pp. 27–32.

[13] B. K. Horn and B. G. Schunck, "Determining optical flow," Artificial intelligence, vol. 17, no. 1-3, pp. 185–203, 1981.

[14] J. Buckley, "Elementary queueing theory based on possibility theory,"

Fuzzy sets and systems, vol. 37, no. 1, pp. 43–52, 1990.

SWITCHED OPTIMAL CONTROL FOR LINEAR SYSTEMS WITH PARAMETRIC VARIATIONS

Salman Zaffar
Department of Electronics and Power Engineering
National University of Sciences and Technology, Pakistan
Email: salmanzafar@pnec.nust.edu.pk

Received: 30-November-2022 / Accepted: 13-December-2022
Karachi Institute of Economics and Technology || Technology Forces Journal, Volume 4, Issue 2, 2022

ABSTRACT

This paper proposes a robust control methodology for dynamical systems with parametric variations. This methodology uses optimal control techniques to propose a controller that guarantees asymptotic stability of such systems. Parametric variations beyond a safe-operation range can cause physical damage to the plant/system. An algorithmic way of determining boundaries of safe-operation with respect to parametric variations is presented. This algorithm proposes a variation-dependent hypersurface that serves as a gain-scheduling criterion for the controller outside the safe-operation range. Bellman's Dynamic Programming is used to construct this algorithm for general dynamical systems. The proposed methodology is applied on an LTI system to show the efficacy of the proposition.

Keywords: Robust Control, Dynamic Programming, Parametric Variation, Optimal Control.

1. INTRODUCTION

Physical properties of dynamical systems are manifested in the form of parameters in their mathematical models. A variation in any physical property is reflected as parametric variation (PV) in mathematical model. All dynamical systems, in reality, may suffer from parametric variations that render their responses unpredictable. These PVs can be classified as either due to external or internal factors. External factors are outside the system such as disturbances and sensor noise. Internal factors are manifested inside the system such as variable poles, zeros, and time constants of the system and frequency dependent uncertainties e.g. unmodeled high frequency dynamics of the system [1]. Therefore, controlling a parametrically variable dynamical system has been an important problem in control engineering.

Optimal control technique such as H_2 cannot guarantee performance and stability of dynamical systems with variable responses. Adaptive control cannot guarantee stability let alone performance in such systems. H_∞ optimal control technique tackles the problem of guaranteeing performance in those systems in which responses vary only due to unmodeled high frequency dynamics (known as unstructured uncertainty). All of these methods fall short of expectation in case of systems which experience real parametric variations. This deficiency leads to robust control methods which have been successful in guaranteeing at least stability of parametrically varying dynamical systems.

A main result in ascertaining the stability of parametrically varying linear systems is that of Kharitonov [2]. Kharitonov's method can be considered as the next-step after Routh-Hurwitz method which is well known in linear dynamical systems.

A few robust control methods relying on optimal control techniques for parametrically varying dynamical systems are given in [3], [4]. In [3], the main proposition is that the parametrically varying states of a nonlinear system can be introduced into a cost functional. It is shown in [3] that the resultant linear quadratic regulator can stabilize the nonlinear system with bounded parametric variations. In [4], a switchable control strategy is presented for guaranteeing stability in parametrically varying linear systems. It is proposed that the control law be switchable so that a controller with more suitable control effort can be selected outside the allowed range of parametric variations. A machine learning algorithm such as support vector machines is used to design a variation dependent hypersurface that serves to determing the size of the parametric variations.

The main technical proposition in this paper is an advancement of the idea presented in our previous work in [4]. The proposed advancement has been carried out with two goals in sight. One is that of traditional stability and the new is that of physical safety. It may be the case that an extraordinary parametric variation outside the safeoperation range is physically damaging for the structure of the plant/system. In such cases, it may be desirable to increase or reduce the control effort outside the safe-operation range so as to minimize the risk of physical damage. For goals such as aforementioned, we propose a switchable control law consisting of a variation-dependent hypersurface as switching criterion. This hypersurface is designed through using notion of Dynamic Programming which was proposed by Bellman in [5]. Rigorous use of dynamic programming leads to the design of hypersurface as shown in this paper. Three ranges of operation are determined for the system in each of which stability and physical safety are design goals. Our proposed approach is general and can be applied to nonlinear systems too; however, for convenience we consider an LTI system to show the efficacy of the proposition.

Rest of the paper is organized as follows: Section II states preliminaries and the problem. Section III presents the LQR derivation through Pontryagin's minimum principle. Section IV analyses the stability of the closed-loop system with the proposed control strategy. Section V presents a motivating example of a double-integrator problem with a single parametric variation. Section VI shows the simulation results. Finally Section VII concludes the paper.

2. PRELIMINARIES & PROBLEM STATEMENT

A. Preliminaries

In simple terms optimal control aims to stabilize a dynamical system at a minimum cost. Dynamical systems can be linear or nonlinear as well as be time-varying or time-invariant. This cost is an index which reflects the performance of the system during any control process such as stabilization or regulation. Cost incurred during a control process is specified in terms of either energies of system variables (such as state and control) or time spent in any control process. Naturally, cost in any case is to be kept at a minimum. We focus on minimizing the energies of system variables in this work.

The concept of cost is mathematically realized as a functional. Functional extremization (minimization or maximization) is common to both the optimal control and its ancestor calculus of variations. Consider a dynamical system along with a quadratic cost functional as shown below

$$\dot{x} = f(x,u) = f_0(x) + g(x)u \tag{1}$$

$$J = \int_{t0}^{T} (x^T Q x + u^T R u)dt + x^T(T) S x(T) \tag{2}$$

where $x \in R^n$, $u \in R^m$ and $[t_0, T]$ is the time duration of interest. Q, R, and S are weighting matrices for state(s), control(s), and terminal value(s) of

state(s), respectively. When the final time T is increased to ∞ meaning that the stabilization of the linear system is relaxed to be carried out without some specific final time, then the cost functional gets modified to the equation given below

$$J = \int_0^{\infty} (x^T Q x + u^T R u) dt \quad (3)$$

and the problem is known as infinite-horizon optimal control problem. It is to be noted that a large r, in case of scalar x and u, relative to q will result in a very restricted control effort allowing for smaller actuators, amplifier gains, and motors to implement a control law. On the other hand, a large q restricts state energy resulting in a raised control effort and damped system.

B. Problem Statement

The desired control law is expected to have switchable capability so as to be able to restrict states and controls within and without safe-operation regions. This capability is to ensure physical safety and possibly stability if not optimality when parametric variations are outside safe-operation range. We propose to modify available optimal controllers to have this switching capability. This will not only ease the controller synthesis but also simplify the stability analysis of the closedloop system.

Now for a dynamical system in (1) with a cost functional in (3), the stabilizing optimal controller will have the following form

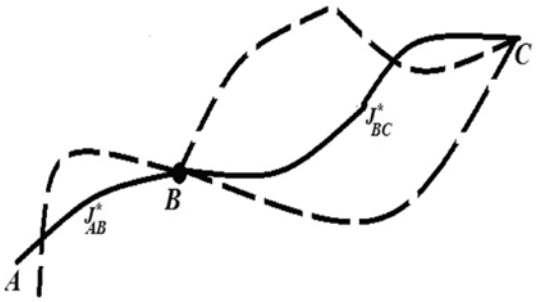


Fig. 1.Graphical representation of Principle of Optimality

$$u = -R^{-1} g^T(x) \frac{\partial J^*}{\partial x} \quad (4)$$

where J^* is cost incurred by system trajectories traversing an optimal path. Since, we focus on minimizing energies in this process hence J^* becomes a positive-definite function. Hence, the resulting control law incorporates *state-feedback* [6].

The switching capability is to be introduced in abovementioned general form of optimal controllers according to following criterion

$$J_i^* = \begin{cases} J_1^* & when\ PV\ is\ above\ safe{-}operation\ range \\ J_0^*\ when\ PV\ is\ inside\ safe-operation\ range \\ J_{-1}^*\ when\ PV\ is\ below\ safe-operation\ range \end{cases} \quad (5)$$

where J_1^* , and J_0^* are J_{-1}^* costs incurred by system trajectories traversing optimal paths in above-, inside-, and below safe-operation range.

3. BELLMAN'S DYNAMIC PROGRAMMING

Bellman's dynamic programming approach is built upon the following Principle of Optimality:

An optimal policy has the property that no matter what the previous decision (i.e., controls) have been, the remaining decisions must constitute an optimal policy with regard to the state resulting from those previous decisions. which is stated mathematically as following

$$J^*(x,t) = \min_{\substack{u(\tau) \\ t \le \tau \le t+\Delta t}} \left[\int_t^{t+\Delta t} L(x,u,\tau) d\tau + J^*(x+\Delta x, t+\Delta t) \right] \quad (6)$$

where $J*(x,t)$ is the optimal cost to go t from time t to time T $\int_t^{t+\Delta t} L(x,u,\tau) d\tau$ is the cost from current time t to time $t+\Delta t$ and $+J^*(x+\Delta x, t+\Delta t)$ is the optimal cost-to-go from time $t+\Delta t$ to time T.

Applying the principle of optimality on dynamical systems in the context of this paper requires i) first order approximation of the system dynamics, and ii) introduction of parametric variations in system model. Therefore, the approximated parameter-dependent model of the system is given below as

$$\dot{x} = f(x,u,v) \Rightarrow \frac{\Delta x}{\Delta t} = f(x,u,v) \Rightarrow f(x,u,v)\Delta t \quad (7)$$

where the factor v represents the parametric variations in

the system.

We will now focus on building the parameter-dependent hypersurface algorithmically. In the first step of this two-step algorithm, we derive the Hamilton-Jacobi-Bellman Equation (HJB) for the system at nominal conditions. In the second step we use this HJB equation to obtain the aforementioned hypersurface.

A. First Step: Solving HJB Equation

Expanding the second term on the right using Taylor's series in (7) and approximating the integral term to first-order yields

$$J^*(x,t) = \min_{\substack{u(\tau) \\ t\le\tau\le t+\Delta t}} \left[L(x,u)\Delta t + J^*(x,t) + (\frac{\partial J^*}{\partial x})^T \Delta x + \frac{\partial J^*}{\partial t}\Delta t\right]$$

$$-\frac{\partial J^*}{\partial t}\Delta t = \min_{\substack{u(\tau) \\ t\le\tau\le t+\Delta t}} \left[L(x,u)\Delta t + (\frac{\partial J^*}{\partial x})^T \Delta x\right]$$

$$-\frac{\partial J^*}{\partial t}\Delta t = \min_{\substack{u(\tau) \\ t\le\tau\le t+\Delta t}} \left[L(x,u)\Delta t + (\frac{\partial J^*}{\partial x})^T f(x,u,v_0)\Delta t\right]$$

now in case of $\Delta t \to 0$

$$-\frac{\partial J^*}{\partial t} = \min_{u(\tau)} \left[L(x,u) + (\frac{\partial J^*}{\partial x})^T f(x,u,v_0)\right] \quad (8)$$

where v_0 in $f(x,u,v_0)$ represents nominal model of system. The last equation is a partial differential equation and is known as Hamilton-Jacobi-Bellmen (HJB) equation. Its solution $J*$ is the optimal cost from the current time t onwards. In case of time-invariant dynamical systems, the equation reduces to

$$0 = \min_{u(t)} \left[L(x,u) + (\frac{\partial J^*}{\partial x})^T f(x,u,v_0)\right] \quad (9)$$

Proceeding with the derivation, if the control input u is explicitly stated outside the system dynamics in above equation, then

$$0 = \min_u \left[L(x,u) + (\frac{\partial J^*}{\partial x})^T (f(x,v_0) + g(x)u)\right] \quad (10)$$

With a quadratic $L(x,u)$, equation given above modifies to

$$0 = \min_u \left[x^T Q x + u^T R u + (\frac{\partial J^*}{\partial x})^T (f(x,v_0) + g(x)u)\right] \quad (11)$$

Hence, the minimizing (optimal) control signal $u*$ which minimizes the HJB equation is given as

$$u^* = -R^{-1} g^T(x) \frac{\partial J^*}{\partial x} \quad (12)$$

For nonlinear and linear dynamical systems, forms of the solutions can be given respectively as

$$J = {}^1\!/_2\, x^T P(x) x, J = {}^1\!/_2\, x^T P x \quad (13)$$

where $P(x)$ is a state-dependent symmetric positive definite matrix and P is a constant symmetric positive-definite matrix. The optimal control signals for nonlinear and linear systems thus become

$$u^* = -R^{-1} g^T(x) P(x) x, \qquad u^* = -R^{-1} B^T P x \quad (14)$$

Inserting these expressions for optimal control into time in variant HJB in (9) results in a State-Dependent-Riccati Equation (SDRE) for nonlinear systems and an AlgebraicRiccati Equation (ARE) for linear systems. An introduction to SDRE is given in [7]. SDRE and ARE can be solved for P(x) and P, respectively. These solution matrices are used in the design of hypersurface as shown ahead. The SDRE and ARE are given below respectively as

$$A^T(x,v_0)P(x) + P(x)A(x,v_0) + Q - P(x)BR^{-1}B^T P(x) = 0 \quad (15)$$

where $A(x,v_0)$ is a special decomposition of $f(x,u,v_0)$,

and

$$A(v_0)^T P + PA(v_0) + Q - PBR^{-1}B^T = 0 \quad (16)$$

where A is a linearized version of $f(x,u,v_0)$.

B. Second Step: Designing Hypersurface

The variation - dependent hypersurface is designed by inserting the solution matrices of SDRE (denoted Ps(x)) and ARE (denoted as Ps) back into the (15) and (16) respectively. The only difference will be that the parametric variations are not fixed in latter case. The expressions for the hypersurfaces for nonlinear and linear systems are

$$Snonlinear: A^T(x,v)P_s(x) + P_s(x)A(x,v) + Q - P_s(x)BR^{-1}B^T P_s(x) \quad (17)$$

$$S_{linear}: A(v)^T P_s + P_s A(v) + Q - P_s B R^{-1} B^T P_s \quad (18)$$

At this point, we focus our attention towards the linear systems and find a general expression for variation-dependent hypersurface through (18). For a two-dimensional linear system with following system matrices

$$A(v) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, Q = \begin{bmatrix} q_{11} & 0 \\ 0 & q_{22} \end{bmatrix},$$
$$R = 1, and\ P = \begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix}$$

The expression for hypersurface of linear system can be expanded using the above mentioned system matrices as shown below

$$\begin{bmatrix} 2ap_1 + 2cp_2 + q_{11} & bp_1 + ap_2 + dp_2 + cp_3 \\ bp_1 + ap_2 + dp_2 + cp_3 & 2bp_2 + 2dp_3 + q_{22} \end{bmatrix} = \begin{bmatrix} p_2^2 & p_2 p_3 \\ p_2 p_3 & p_3^2 \end{bmatrix} \quad (19)$$

The desired hypersurface can be chosen from any of the three equations given above. Therefore, the three hypersurfaces can be

$$H_1: 2ap_1 + 2cp_2 + q_{11} - p_2^2, \text{ or} \quad (20)$$

$$H_2: bp_1 + ap_2 + dp_2 + cp_3 - p_2 p_3, \text{ or} \quad (21)$$

$$H_3: 2bp_2 + 2dp_3 + q_{22} - p_3^2 \quad (22)$$

These hypersurfaces can be used to form a switching function. We proposed one similar function in our previous work [4] where switching took place between a couple of controllers. Each of these hypersurfaces can be used to obtain a switching function that can allow for switching between infinite number of controllers. In present case, we consider switching only between three controllers. The switching function is given as

$$f(H,Y) = \frac{1}{2}\{sgn(H - X_1) + sgn(H + X_2)\} \quad (23)$$

where X_1 and X_2 are maximum and minimum values of chosen hypersurface at most and least allowable values of parametric variations. This means that X_1 = max(H) and X_2 = min(H). These values represent safe-operation boundaries above and below the hypersurface. This switching function has a ternary output i.e. f(H,X) : Rn ⇒ {+1,0,−1} which is related to the PVs as follows

$$f(H,Y) = \begin{cases} +1 & when\ H \geq X_1 \\ 0 & when\ X_2 \leq H \leq X_1 \\ -1 & when\ H \leq X_2 \end{cases} \quad (24)$$

Design of hypersurface and switching function leads to three ranges of operation. For each of these ranges, a different ARE can be solved to select suitable gains. Hence, the switchable control signal is given as below

$$u = -R^{-1} B^T (P^s) \quad (25)$$

where

$$P^s = \begin{cases} P^{X_1} & when\ f(H,X) \geq +1 \\ P & when\ f(H,X) \geq 0 \\ P^{X_2} & when\ f(H,X) \geq -1 \end{cases} \quad (26)$$

where P^{X1}, P, and P^{X2} are respective solutions of AREs solved for above safe-operation range, inside safe-operation range, and below safe-operation range.

4. STABILITY ANALYSIS

Stability of switched systems is difficult to determine through analytical methods. Since we employ switching to switch between three controllers, following result is helpful in evaluating stability of in our case. This result says that for two LTI subsystems, U and Y, a necessary and sufficient condition for the existence of a common quadratic Lyapunov function (CQLF) is that the matrix products $U \times Y$ and $U \times Y^{-1}$ do not have any real negative eigenvalues. Proof that a CQLF guarantees stability two switched LTI system is given in [8]. We choose two closed-loop systems, at a time, formed from the above three ranges of operation and apply the result and find that the switched system is stable.

5. MOTIVATING EXAMPLE

Lets consider the following system matrices for the double integrator problem,

$$A(v) = \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix}$$

1) Solving ARE: The ARE is solved for the system having following matrices

$$A(v_0) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$
$$R = 1, and\ P = \begin{bmatrix} 1 & 1.414 \\ 1.414 & 1 \end{bmatrix}$$

and thus the optimal control law becomes

$$u = -p_2x_1 - p_3x_2 = -x_1 - 1.414x_2 \quad (27)$$

2) Designing Hypersurface: Manipulating (18) for double-integrator problem yields the following

$$\begin{bmatrix} q_{11} & bp_1 \\ bp_1 & 2bp_2 + q_{22} \end{bmatrix} = \begin{bmatrix} p_2^2 & p_2p_3 \\ p_2p_3 & p_3^2 \end{bmatrix} \quad (28)$$

Choosing the following hypersurface

$$H: 2bp_2 + q_{22} - p_3^2 = 2b - 1 \quad (29)$$

After having formed the hypersurface, boundary for switching can be found by assuming the allowable upper and lower values of the b to be 1.25 and 0.75 respectively. So when b = 1.25 the non-zero scalar value of X_1 is 1.5 and when b = 0.75 the non-zero scalar value of X_2 is 0.5. Therefore, the switching function is

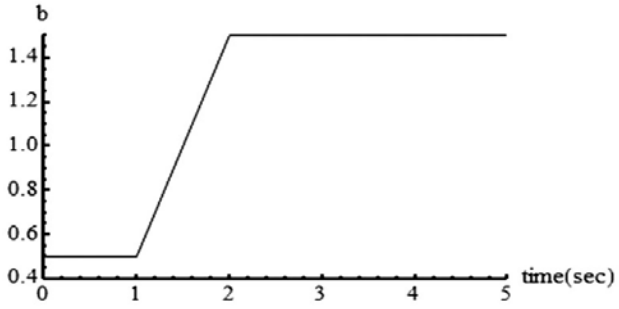


Fig. 2. Parameter b's variation from a value of 0.5 to a value of 1.5.

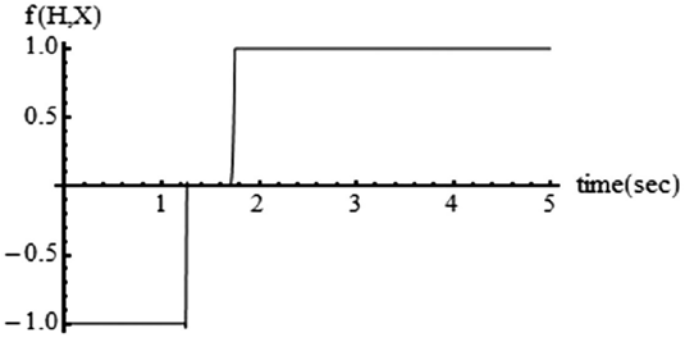


Fig. 3. Switching function producing +1,0,−1 with respect to parametric variation in b.

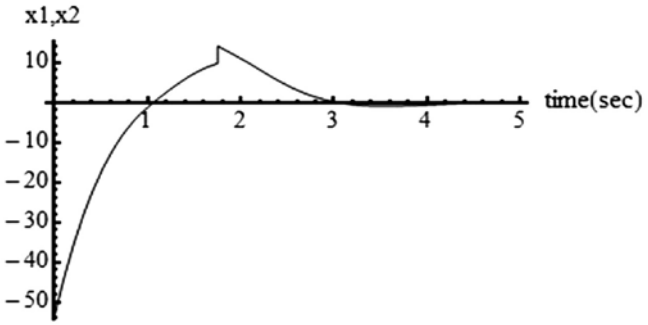


Fig. 4. Switched control signal exhibiting switchings

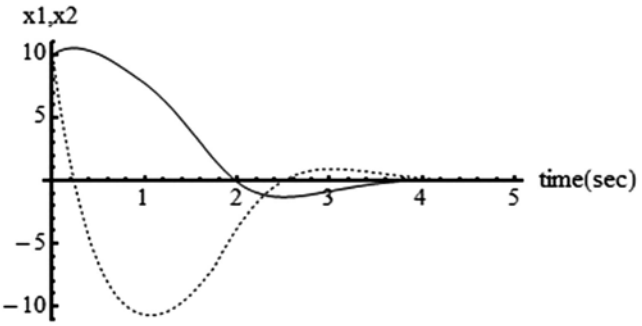


Fig. 5. Asymptotically stable states x1(solid) and x2(dashed) converging to

x =0

$$f(H,Y) = \frac{1}{2}\{sgn(H - 1.5) + sgn(H + 0.5)\} \quad (30)$$

Now AREs are solved to obtain LQR gains for above and below the safe-operation range. The system matrices for these two range are

$$A^{X1} = \begin{bmatrix} 0 & 1.25 \\ 0 & 0 \end{bmatrix}, Q^{X1} = \begin{bmatrix} 10 & 0 \\ 0 & 0 \end{bmatrix}, and$$
$$A^{X2} = \begin{bmatrix} 0 & 0.75 \\ 0 & 0 \end{bmatrix}, Q^{X2} = \begin{bmatrix} 10 & 0 \\ 0 & 0 \end{bmatrix}, yielding$$
$$P^{X1} = \begin{bmatrix} 7.11 & 3.16 \\ 3.16 & 2.81 \end{bmatrix}, P^{X2} = \begin{bmatrix} 9.18 & 3.16 \\ 3.16 & 2.17 \end{bmatrix}$$

respectively.

The stability analysis of the switched system is now carried out. Two controllers at a time are chosen and tested for stability using the result in [8]. The three closed-loop systems formed for three ranges of operation are given as

$$A - B.(R^{-1}B^TP) = \begin{bmatrix} 0 & 1 \\ -1 & -1.414 \end{bmatrix}$$
$$A^{X1} - B.(R^{-1}B^TP^{X1}) = \begin{bmatrix} 0 & 1.25 \\ -3.16 & -2.81 \end{bmatrix}$$
$$A^{X2} - B.(R^{-1}B^TP^{X2}) = \begin{bmatrix} 0 & 0.75 \\ -3.16 & -2.17 \end{bmatrix}$$

Any combination of two closed-loop systems out of the three satisfy the result for asymptotic stability of switched systems. Therefore, a CQLF exists guaranteeing the asymptotic stability of the switching.

6. SIMULATIONS

Figure 2 shows the parametric variation in *b*. This parameter varies from is allowed to vary in the range from 0.5 to 1.5. Parametric variation in this range activate all three ranges of operation. Figure 3 shows the switching function that outputs +1,0,−1 with respect to the parametric variation *b*. Figure 4 shows the switched control signal. Finally, figure 5 shows the state convergences to x = 0.

7. CONCLUSIONS

This paper has proposed a robust control methodology that guarantees stability of dynamical system in presence of parametric variations. Three ranges of operation are identified on the basis of values of these variations. It is shown that the controller can ensure stability and physical safety of systems in all the three ranges of operation. The mathematical treatment is general enough to be applicable to nonlinear systems as well. Future work can include a design of an observer and a stability analysis method in case of nonlinear systems.

REFERENCES

[1] Skogestad S, Multivariable Feedback Control. England: Wiley; 2005.

[2] S.P. Bhattacharyya, H. Chapellat and L. Keel, Robust Control: The Parametric Approach. USA: Prentice-Hall; 1995.

[3] Feng Lin; Brandt, R.D., "An optimal control approach to robust control of robot manipulators," in Robotics and Automation, IEEE Transactions on , vol.14, no.1, pp.69-77, Feb 1998

[4] Zaffar, S.; Memon, A.Y., "Robust and optimal stabilization of uncertain linear systems using LQR methods," in Control (CONTROL), 2014 UKACC International Conference on , vol., no., pp.163-167, 9-11 July 2014

[5] R. E. Bellman. Dynamic Programming. Princeton University Press, Princeton, N J, 1957.

[6] M. Athans and P. Falb. Optima Control: An Introduction to The Theory and Its Applications. McGraw Hill Book Company, New York, NY, 1966.

[7] Cloutier, James R., "State-dependent Riccati equation techniques: an overview," in American Control Conference, 1997. Proceedings of the 1997 , vol.2, no., pp.932-936 vol.2, 4-6 Jun 1997

[8] Shorten, R. N. and Narendra, K. S., Necessary and sufficient conditions for the existence of a CQLF for a finite number of stable LTI systems,International Journal of Adaptive Control and Signal Processing, Volume 16, Number 10, pp. 709-728, December 2002

KARACHI INSTITUTE OF ECONOMICS AND TECHNOLOGY
Main Campus, PAF Airmen Academy, Korangi Creek
Tel: 021-35091114 - 7 | Cell: 0345-4982304-7

Designed & Produced by HAWK