Comparison of User Satisfaction between Text and Virtual Reality Based Expert Systems for Diagnosing Cattle Diseases

Kabo Nkabiti, Malatsi Galani, and Adamu Murtala Zungeru

***Abstract*—Even though usability testing of software applications has been studied for decades, little empirical researches have been conducted in Agriculture domain, especially in comparing the usability of none- immersive virtual reality and text based diagnosis expert systems for diagnosing cattle diseases. A quasi -experimental research was conducted in which 62 cattle farmers both commercial and subsistent used either a non-virtual reality based (System A) or a text-based (System B) expert systems to diagnose cattle diseases. An attitude rating scale Software Usability Measurement Inventory (SUMI) was used to collect data and the reactions towards the expert-based systems. The measures were efficiency, the affect, helpfulness of the system, controllability system, and system Learnability. The results obtained from the research indicated that the group of farmers using System A was more satisfied as compared to those using system B.**

***Index Terms*—Usability, human computer interaction, non-immersive virtual reality, expert system, cattle diseases, user satisfaction.**

I. Introduction

To set out an atmosphere for this research paper we start by giving a brief overview of usability. Usability is a key issue in human-computer interaction (HCI) since it is the aspect that commonly refers to the quality of the user interface [1], Usability is defined by International Standards Organization 9241-11 as [2] "The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context." This study is focused on user satisfaction which is defined as the perceived usability of the overall system, and how its users accept it and the people affected by the system. Satisfaction may be measure on specific aspects of the system or may be used to measure the overall system [3]. It is essential for software, websites, mobile application, or any user-operated technology developers and designers to create systems which satisfy the needs of the users [4].Despite an increased interest in usability testing of software applications, it is surprising that so little empirical researches have been conducted in this area of study, especially in comparing the usability of non-immersive virtual reality and text based diagnosis expert systems for cattle diseases. In this study, the focus is to compare the user satisfaction of the two diagnosis expert systems for cattle diseases. Different expert systems for diagnosing diseases have been developed mostly using CLIPS shell [1]-[3] which are text based. Some of these systems use software such as ESTA [4]-[6] which is an "expert system shell developed by Prolog that allows the use of using pictures in the title, parameter, in advice and in action list" [7]-[9].

We experimented with two expert systems that use different user interfaces to determine user's satisfaction. The two systems compared were text-based expert systems (System B) developed using CLIPS, expert system development shell, and the virtual reality based expert systems developed using 3ds max 2016. We added interactivity to the 3D models and animation using Unity software. This helped to show disease symptoms and signs more clearly.

In our investigation, we aimed to address the following research question;

- Is there a significant difference in user satisfaction between farmers using text and non-immersive virtual reality based cattle diseases diagnosis expert system?

The paper is arranged and divided into the following sections, firstly we have the introduction which introduces the problem and gives an insight into user satisfaction. The second section is the related work which looks at what has already been done which is similar to our research. Under section three we also discuss the methods used to archive the main of the study, the participants, and the procedure. The fourth section is the result presentation. The fifth section is results discussion which includes its limitations, implications and comparison to the existing results then the last section is the conclusions.

II. Related Work

This sub-section describes similar works and discusses the results obtained from those studies.

Ogogzake [10] carried an experimental research by comparing the use text and multimedia interface for the provision of information about drugs prescription to old people with an average of 71 years. The participants (62 divided into two groups) compared the paper

Manuscript received October 15, 2016; revised January 6, 2017. This work was supported and sponsored by management of Botswana International University of Science and Technology in partnership with Department of Computer Science and Information Systems and Department of Electrical, Computer and Telecommunication Engineering.

K. P. Nkabiti and M. Galani are with the Department of Computer Science and Information Systems, Botswana International University of Science and Technology, Palapye, Botswana (e-mail: kabo.nkabiti@studentmail.biust.ac.bw, galanim@biust.ac.bw).

A. M. Zungeru is with the Department of Electrical, Computer and Telecommunication Engineering, Botswana International University of Science and Technology, Palapye, Botswana (e-mail: zungerum@biust.ac.bw).

doi: 10.18178/ijiee.2017.7.5.678

information-based system with a computerized information system. The results indicated that the group of users using multimedia-based presentations got much better results, the users preferred to use multimedia interfaces. The electronic system performed better that the text based. In terms of gender, men preferred both the multimedia and text-based interfaces whereas women found the multimedia system more pleasant than the text-based interface.

In another study, two teaching methods for presentation of learning materials to pediatric nursing students were compared by Granados [9] at the University of San Francisco. Text-based modules and multimedia based modules were compared. The independent variable of the research was "knowledge acquisition of mathematical calculation skills for medical administration" [11]. The research defined calculation skills as the ability of the student to calculate (1) weight–based safe dose ranges, (2) intravenous flow rates for primary and secondary medical infusion, (3) Converting pounds to kilograms, and lastly (4) the maintenance of fluid [5]. The results of the research indicated that there was no statistical significant between the two modules in respect of the pass rate, as well the four items (1), (2), (3) and (4) mentioned above. There was also an indication of the results that the modules were not as effective as anticipated for teaching and delivering the pediatric medical administration materials. In addition to the findings, it was also evident that the preponderance of errors was similar for both groups with an exception of only a few errors with a minor statistical significant.

Also, Jung-Wei Chen *et al*. [12] conducted a research on a comparison of graphical user interface (commonly considered to be better) with text-based interfaces for a dental records system. Two groups of users were evaluated in this study, the novice, and the expert users to check the time and steps taken to complete the tasks. Different usability evaluation and analysis techniques such as task analysis, GOMS (Goals, Operations, Methods and Selection rules), HTA (Hierarchical task analysis) and Distributed representation analysis were used to compare the effectiveness of the systems. To assess whether experience affects the task performance of the systems the within-subject design evaluation method was used. The results indicated that for experts' user the graphical user interface (GUI) was not better that the text-based interface (TBI). The novice users using the systems with GUI performed better than those using the TBI. The research concluded by indicating that the user-friendliness of the user interface of the system depends on how the task was mapped which means that GUI can be better than TBI or may not be better it only depends on the task being carried out.

Furthermore in 1996 Ramsey [13] conducted a research on a comparison of three instructional learning methods which includes a text-based instructional system and two multimedia systems. The two systems using multimedia instructions used different interfaces the topic-oriented and problem-solving interface. Intellectual skills and verbal information were the domains of learning being examined in this experimental research design. The two designs used for examination of original learning as well as the retention components of learning [13]. The research measured the original learning first then followed by instructional treatment. Then after a fortnight the retention will be measured. There were 30 participants participating in each session. Analysis of variance was used to analyze the scores. The results obtained indicated that there was no statistical significance difference between the text-based and the two multimedia systems during both domains original learning and retention sessions [13].

Lastly, Cheon and Grant [14] compared the learning efficiency of different types of interfaces which were based on the cognitive load theory. The interfaces being compare included the text-based, graphical interface as well as the metaphorical interfaces so as to establish which interface has a lower cognitive workload. Fifty undergraduates metal effort levels and their performance score we compared. The results obtained indicated that there was no statistical significance difference between the groups in terms learning efficiency.

From the reviewed related work it evidently showed that studies on a comparison of user satisfaction between text-based and non-immersive virtual reality based expert systems are very limited. Non-immersive virtual reality application on animal health expert systems is a new concept so there is limited research made on the field [15], [16]. With the factors identified above the researcher found it fit to do research on this title and hopefully this paper will contribute knowledge in this area of study.

There are two factors which distinct our research from the related works. Firstly the comparison that we conduct is based upon text based and virtual reality (interactive multimedia) based expert system, unlike the most related works which compared the text with normal multimedia (images) [5], [10], [12], [13]. Secondly, we adopted the software usability measurement inventory questionnaire (SUMI) to assess the participant's satisfaction with both systems and evaluate their preference. The ISO 9241 standard regards SUMI as a method used to test the satisfaction of a user [2], [17]. This common usability tool is made up of validated 50-questions in which participants score each question on a scale of three points which are agree, undecided or disagree. The participants were allowed to ask the researcher questions during task performance. The questionnaire (SUMI) was answered by the participants immediately after using the systems. The data collected from the questionnaire was then sent to Dr. Jurek Kirakowski (SUMI employee) for analysis. The researcher managed to acquire the academic SUMI license which is free [18]-[20].

III. Measures of User Satisfaction

Different studies have indicated that user satisfaction has five measures a listed below;

- Efficiency: This is how the user feels about systems responds when tasks are performed through it which includes aspects such as time taken to complete a task, resources used to complete the task by means of the systems performance [11].
- Affect: This is referred to as emotional feeling in psychology. In the context of our case, it means when the

participant is feeling mentally stimulated in an unpleasant way or in a pleasant way after using the systems [18].

- Helpfulness: This whereby the user perceives the system to be communicating in a helpful manner, and there is a proper assistance to resolve the problems encountered while operating the system [18].
- Control: This is the degree at which the user feels that he/she has power over the system and he/she, not the system is pace setter [18].
- Learnability: the ability of the user to learn new features of the systems at ease and get easily gets started.

IV. The Hypothesis

The hypotheses that were tested in the experiment were as follows;

Hypothesis 1: There is no significant difference in the type of cattle disease diagnosis expert system text or virtual reality based on tasks efficiency.

Hypothesis 2: There is no significant difference in the type of cattle disease diagnosis expert system text or virtual reality based on tasks affect.

Hypothesis 3: There is no significant difference in the type of cattle disease diagnosis expert system text or virtual reality based on tasks helpfulness.

Hypothesis 4: There is no significant difference in the type of cattle disease diagnosis expert system text or virtual reality based on tasks controls.

Hypothesis 5: There is no significant difference in the type of cattle disease diagnosis expert system text or virtual reality based on tasks learnability.

Hypothesis 6: There is no significant difference in the type of cattle disease diagnosis expert system text or virtual reality based on tasks in general.

V. Methods

This section provides information for judging the validity of this study. It is a clear description of how the experiment for this study was done.

A. Quasi-Experimental Design

The independent variables tested were the usability measures task efficiency, task affect, task control, task learnability and task helpfulness tested on both text and virtual reality based expert systems. The between-subject design was used to compare the text and the virtual reality based expert systems. The 62 participants were dived into half and the half used text based system whereas the other half used the virtual reality based expert systems. Both interfaces were desktop based. Usability testing software called Morae from TechSmith the use of the systems by the participants [21], [22]

B. Participants

We tested the two systems with 64 individuals and they were divided into a group of 32 participants each. The groups included both commercial and subsistent farmer young and old mostly those who could and read and were registered with the department of veterinary services which is where they were recruited. The age range of the participants was approximately between 18 and 65. The participants were assured that the information they shared with us would be kept private and confidential and it would be anonymous [23], [24]. Procedure

When starting up the experiment the participants were giving a short lesson on how the systems works and they familiarized themselves with its purpose in general. They were informed that they have limited time to complete different tasks. For the situation to be realistic the participants were requested to imagine diagnosing their cattle in real life as opposed to just using their knowledge to as or even their opinion when performing the set tasks.

As the participants performed the given tasks they were left alone in a room and Morae recorder recorded a video of everything they were doing. After the participants completed all the tasks they signaled the experimenter and he gave them a SUMI questionnaire to complete for user satisfaction assessment which determined their preferred system.

The procedures for both systems were the same. Even some participants had difficulty using the systems the experimenter gave them more time to familiarize themselves with the systems, but otherwise, most participants did not have any difficulties understating how the systems work.

VI. SUMI Results

Fig. 1 below shows the SUMI scale profiles for the means and standard deviations for the VR-based expert system. According to Erik and Veenendaal [25] "The SUMI database states that 50 is the average value for the global (General) score in a normal distribution and the standard deviation is 10". When a value is greater than 50 then user satisfaction is considered to be more that average. The data from Fig. 1, 2, 3, 4 were used for statistical assessment for comparing the two systems.

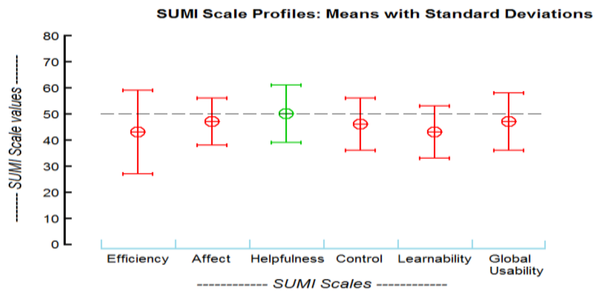

Fig. 1. VR means and SD.

	Mean	St Dev	Median
Global	46.81	11.12	43.5
Efficiency	43.19	15.58	39.0
Affect	46.56	9.23	45.0
Helpfulness	49.66	11.32	47.5
Controlability	46.34	10.48	44.0
Learnability	42.78	10.47	41.0

Fig. 2. Virtual reality-based system means and SD and median.

An overview of means, p-values, and independent-samples t-test for the total scores of user satisfaction (efficiency, affect, helpfulness, controls, learnability and systems in general) for both virtual reality and text based expert systems are presented in tables below;

TABLE I: EFFICIENCY

Compared Systems	Means	P-value(0.05)	Hypothesis 1
System A (VR)	38.0935	0.0010	Rejected
System B (Text)	43.1875		

Because the sample size for both systems was the same there was an assumption for equal variances. The results of an independent-samples t-test for Table I (Efficiency) confirm that there is a significant difference between Systems A and System B because p-value is 0.0010 which is less than 0.05. The mean difference shows that System A (VR) has a lesser mean than System B (Text) which makes System A (VR) more efficient than System B (Text).

TABLE II: AFFECT

Compared Systems	Means	P-value(0.05)	Hypothesis 2
System A (VR)	44.4688	0.1594	Accepted
System B (Text)	46.5625		

With regard to affect the results in Table II indicates that there is no significant statistical difference between the two group's affect scores. The results were also acquired through independent-samples t-test, therefore, there was no effect size computed for the total affect to determine which is better.

TABLE III: HELPULNESS

Compared Systems	Means	P-value(0.05)	Hypothesis 3
System A (VR)	41.25	0.0001	Rejected
System B (Text)	49.6562		

As for the systems, Helpfulness the independent-samples t-test was used to examine whether there is any statistical difference between the groups. The results indicated that there is an enormous statistical difference between the groups as the p-value (0.0001) was way less than 0.05. The difference of the means was also huge with System A (VR) have a smaller mean making it more helpful as compared to System B (Text).

TABLE IV: CONTROL

Compared Systems	Means	P-value(0.05)	Hypothesis 4
System A (VR)	43.2815	0.0416	Rejected
System B (Text)	46.3435		

With regard to control, an independent-samples t-test was also conducted to examine if there is any significant statistical difference. The results indicate that there is a slight statistical significant difference between the groups as the p-value (0.0416) is less than 0.05. System A (VR) is slightly more controllable because its mean is less than that of system B (Text).

TABLE V: LEARNABILITY

Compared Systems	Means	P-value(0.05)	Hypothesis 5
System A (VR)	41.8437	0.5256	Rejected
System B (Text)	42.7812		

There was no significant statistical difference between the groups in learnability as indicated by Table V. The results obtained were computed through the independent-samples t-test. For that reason, the effect size was not computed for learnability means.

TABLE VI: GENERAL

Compared Systems	Means	P-value(0.05)	Hypothesis 6
System A (VR)	41.4375	0.0006	Accepted
System B (Text)	46.8125		

Lastly, the entire system (Global/General) was tested through independent-samples t-test and the results indicated that there was a big difference in reactions to the two systems in general. Statistically, there was a significant difference between the systems. But System A (VR) performed better than System B (Text) because it had a smaller mean number. This means that System A (VR) is satisfied the users much better than the system.

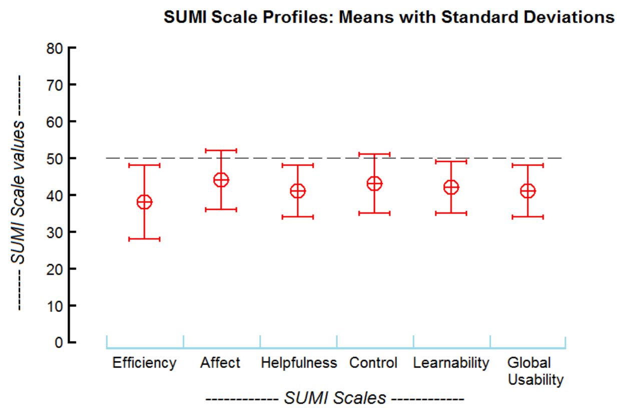

Fig. 3. Text-based system means and SD.

	Mean	St Dev	Median
Global	41.44	7.37	40.0
Efficiency	38.09	9.69	37.0
Affect	44.47	7.84	44.5
Helpfulness	41.25	7.26	42.5
Controlability	43.28	7.58	43.0
Learnability	41.84	7.34	40.0

Fig. 4. Text-based system means and SD and median results presentations.

VII. DISCUSSION OF RESULTS

On this section, a summary of findings from this study is presented. The contribution of this research and its implication on the future work is also presented. Lastly, we conclude with the limitations and recommendations for feather research.

A. Results Implications

The results acquired in this study implies that non-immersive virtual reality expert systems used for cattle disease diagnosis are more efficient, easy to control, easy to learn, helpful and more affect when compared to a text-based expert system. The results uncovered that the users preferred using an interactive multimedia-based system over a text only based system. Interactive multimedia-based system seemed to provide characteristics of a good interface such as clarity, concise, responsive (even though not faster than a text-based one), consistency and attractive.

B. Existing Results Comparison with Our Work

Even though a number of studies have been conducted comparing text and multimedia based systems, most of them were focused on images, only a few covered interactive multimedia particularly virtual reality or augmented reality. It has evidently shown that there is little to no published scholarly articles that compare of user satisfaction between text and virtual reality based expert systems for diagnosing cattle diseases.

C. Results Limitations

Due to limited time and resources a convenient smaller sample size of 32 in each group was used [26]. There this could have could lead to a poor interpretation of results more especially the confidence level and the p-value. None the less small sample size enabled us to address the research questions in a shorter period of time. The other problem we faced was our target ordinance some of our participants did not take the study seriously. The other problem we encountered was that we were looking at the entire cattle farmers' regardless of whether one is a commercial, subsistent or aspiring and the challenge was that they were faced with different problems.

VIII. Conclusions

The overall analysis as indicated by the general or global aspects showed that system A performed much better than system B even though two aspects affect and learnability did not have any significant statistical difference. Most of the participants were very impressed with non-immersive virtual reality technology being used in agriculture for cattle diseases diagnosis. They thought the adoption of such a system could help them manage their farmers well and educate them on different cattle diseases which in return can save money and time if utilizing such a system. A few of the participants, however, showed displeasure in the perception that computers and Smartphone's are for people residing in urban areas and that other farmers might have inconveniences getting used to the new innovation.

The study also has shown that it is not easy to make conclusions about cattle farmers more especially the subsistent farmer because it consists of the population that different and heterogeneous. Nevertheless, we hope that this study will have a contribution to the body of knowledge of interface designs for expert systems in agriculture more especially cattle farming. We intend to expand our study in the future and compare three interfaces for cattle diseases diagnostic expert systems which consist of two multimedia based interfaces (non-immersive virtual reality and augmented reality) with Text using mobile platforms. We would like to examine the preference of the three interfaces in terms of gender and age.

Acknowledgment

The study was completed as one of the objectives of an MSc thesis at Botswana International university of Science and technology who sponsored in all the endeavors of this research study. Sincere gratitude also goes to the Department of Veterinary Services (DVS) who allowed me to use their clients as my participants. Dr. Galani Malatsi also helped in everything he could in this study. Dr. Jurek Kirakowsk and the SUMI team also helped me a lot with the results. Lastly, I would like to thank my colleagues for the support they gave me.

References

[1] S. Parlangeli, E. Marchingiani, and Bagnara, "Multimedia in distance education: Effects of usability on learning," *Interacting with Computers*, vol. 1, no. 12, pp. 37–49, 1999.

[2] N. Bevan, "International standards for HCI," *Encycl. Hum. Comput. Interact.*, vol. 55, no. May, pp. 1–15, 2006.

[3] N. Bevan and M. Macleod, "Usability measurement in context," *Behav. Inf. Technol.*, vol. 13, no. 1–2, pp. 132–145, 1994.

[4] L. Hasan, "Evaluating the usability of educational websites based on students' preferences of design characteristics," vol. 3, no. 3, pp. 179–193, 2014.

[5] M. Hazman, A. Eldin, and A. Yassin, "Animal knowledge based systems in Egypt," vol. 71, no. 21, pp. 33–38, 2013.

[6] S. S. A. Naser, A. Z. A. Ola, and I. Technology, "An expert system for diagnosing eye Diseases," pp. 923–930, 2008.

[7] P. M. N. Rani, T. Rajesh, and R. Saravanan, "Expert systems in agriculture : A review," vol. 3, no. 1, pp. 59–71, 2011.

[8] S. K. Jha, "Development of knowledge base expert system for natural treatment of diabetes disease," vol. 3, no. 3, pp. 44–47, 2012.

[9] R. Prasad, K. Rajeev, and A. K. Sinha, "AMRAPALIKA : An expert system for the diagnosis of pests, diseases, and disorders in Indian mango," vol. 19, pp. 9–21, 2006.

[10] V. Z. Ogozalek, "A comparison of the use of text and multimedia interfaces to provide information to the elderly," pp. 65–71, 1994.

[11] R. Granados, "A comparison of two teaching methods for pediatric medication administration : Multimedia and text-based modules," 2013.

[12] J. Chen and J. Zhang, "Comparing text-based and graphic user interfaces for novice and expert users," pp. 125–129, 2007.

[13] T. D. Ramsey, "The effects of multimedia interface design on original learning and retention," December, p. 107, 1996.

[14] J. Cheon and M. Grant, "Do flashy interfaces help ? Comparing learning efficiency of different interface types based on cognitive load theory," pp. 2003–2007, 2003.

[15] M. Haubner, C. Krapichler, L. Andreas, K. Englmeier, and W. Van Eimeren, "Virtual reality in medicine – Computer graphics and interaction techniques," vol. 1, no. 1, pp. 61–72, 1997.

[16] R. Albert *et al.*, "Virtual reality as a tool for delivering ptsd exposure therapy," pp. 1–22, 2005.

[17] N. Bevan, "Measuring usability as quality of use," vol. 150, 1995.

[18] T. Arh *et al.*, "A case study of usability testing – The SUMI evaluation approach of the EducaNext portal," vol. 5, no. 2, 2008.

[19] Z. Mansor, Z. M. Kasirun, S. Yahya, and N. H. Arshad, "The evaluation of webcost using software usability measurement inventory (SUMI)," vol. 2, no. 2, pp. 197–201, 2012.

[20] J. A. Ogolla, "Usability evaluation : Tasks susceptible to concurrent think-aloud protocol master thesis-TQDV30."

[21] TechSmith, "Morae software in the usability testing lab," Michigan, 2013.

[22] TechSmith, "Software usability testing with morae," Michigan, 2009.

[23] G. Crow and R. Wiles, "Managing anonymity and confidentiality in social research: The case of visual data in Community research."
[24] L. Ted and P. John, "Protecting research confidentiality: Towards a," vol. 21, no. 1, 2006.
[25] D. Erik and P. W. M. V. V. Cisa, "Questionnaire based usability testing," November, pp. 1–9, 1998.
[26] R. V Lenth, "Some practical guidelines for effective sample-size determination," pp. 1–11, 2001.

Kabo P. Nkabiti received the MSc in information systems from Botswana International University of Science and Technology, the BSc Hons in computing (computer science) from Open University UK corresponding with Botho University in Botswana. He is currently a teaching assistant at Botswana International University of Science and Technology. His research interests are virtual reality, human computer integration, artificial intelligence (AI), computer architecture & engineering (ARC), bio-systems & computational biology (BIO), human-computer interaction (HCI), unmanned aerial vehicle (UAV) precision agriculture and wireless sensor networks. He is aspiring to further his studies in PhD in the near future.

Malatsi Galani has a PhD from the University of Wollongong's (Australia) School of Computing and Information Technology. His thesis was titled "Factors affecting nursing staff use of nursing information systems in residential aged care homes", and also has an advanced master's degree in ICT from the same institution. Further, he has bachelor of education (computer science) from the University of Botswana. His research interests are cognition, human computer interaction, information science, educational technology, healthcare informatics, curriculum theory and tele health. He has authored and published a number of papers in journals and international refereed conferences.

Adamu Murtala Zungeru received the PhD, MSc and BEng from Nottingham University, Ahmadu Bello University Zaria Nigeria and Federal University of Technology Minna Nigeria respectively. He was a research fellow in the Electrical Engineering and Computer Science Department at Massachusetts Institute of Technology (MIT) USA, where he also obtained a Postgraduate Teaching Certificate in 2014. He is currently a senior lecturer at Botswana International University of Science and Technology (BIUST). His research interests are in swarm intelligence, wireless sensor networks, embedded systems, and design of analog and digital electronic circuits. He has authored three academic books and has published over forty papers in journals and international refereed conferences. He has also filed 2 international patents application for his innovative work under the Patent Cooperation Treaty (PCT). He is a member of the IEEE, ACM and AP-S.