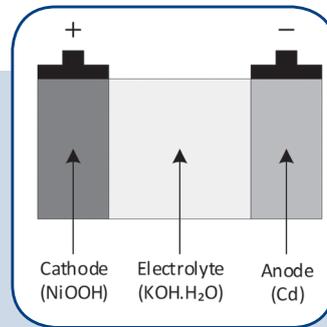
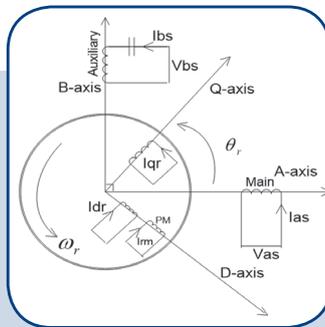
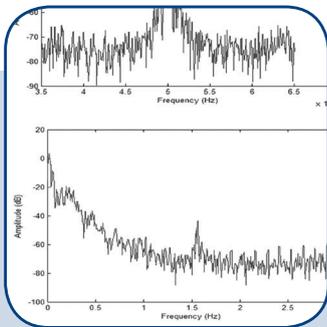
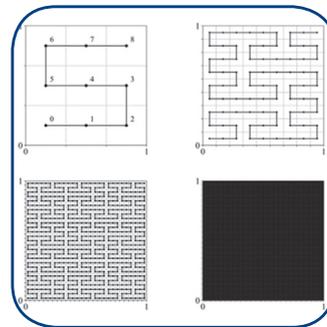
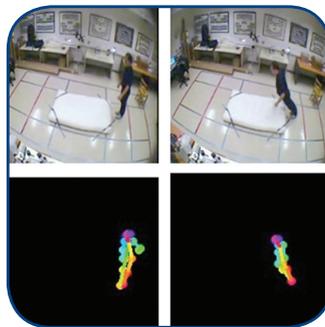


International Journal of Electrical and Computer Engineering Systems



INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

Published by Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia

Osijek, Croatia | Volume 12, Number 1, 2021 | Pages 1-65

CONTACT

International Journal of Electrical and Computer Engineering Systems (IJECES)

Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b, 31000 Osijek, Croatia
Phone: +38531224600, Fax: +38531224605
e-mail: ijeces@ferit.hr

Subscription Information

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.
Giro account: 2390001 - 1100016777,
Croatian Postal Bank

EDITOR-IN-CHIEF

Drago Žagar

J.J. Strossmayer University of Osijek,
Croatia

MANAGING EDITOR

Goran Martinović

J.J. Strossmayer University of Osijek,
Croatia

EXECUTIVE EDITOR

Mario Vranješ

J.J. Strossmayer University of Osijek, Croatia

ASSOCIATE EDITORS

Krešimir Fekete

J.J. Strossmayer University of Osijek, Croatia

Damir Filko

J.J. Strossmayer University of Osijek, Croatia

Davor Vinko

J.J. Strossmayer University of Osijek, Croatia

EDITORIAL BOARD

Marinko Barukčić

J.J. Strossmayer University of Osijek, Croatia

Leo Budin

University of Zagreb, Croatia

Matjaz Colnarič

University of Maribor, Slovenia

Aura Conci

Fluminense Federal University, Brazil

Bojan Čukić

West Virginia University, USA

Radu Dobrin

Mälardalen University, Sweden

Irena Galić

J.J. Strossmayer University of Osijek, Croatia

Radoslav Galić

J.J. Strossmayer University of Osijek, Croatia

Ratko Grbić

J.J. Strossmayer University of Osijek, Croatia

Marijan Herceg

J.J. Strossmayer University of Osijek, Croatia

Darko Huljenić

Ericsson Nikola Tesla, Croatia

Željko Hocenski

J.J. Strossmayer University of Osijek, Croatia

Gordan Ježić

University of Zagreb, Croatia

Dražan Kozak

J.J. Strossmayer University of Osijek, Croatia

Sven Lončarić

University of Zagreb, Croatia

Tomislav Kilić

University of Split, Croatia

Ivan Maršić

Rutgers, The State University of New Jersey, USA

Kruno Miličević

J.J. Strossmayer University of Osijek, Croatia

Tomislav Mrčela

J.J. Strossmayer University of Osijek, Croatia

Srete Nikolovski

J.J. Strossmayer University of Osijek, Croatia

Davor Pavuna

Ecole Polytechnique Fédérale de
Lausanne, Switzerland

Nedjeljko Perić

University of Zagreb, Croatia

Marjan Popov

Delft University, The Netherlands

Sasikumar Punnekkat

Mälardalen University, Sweden

Chiara Ravasio

University of Bergamo, Italy

Snježana Rimac-Drlje

J.J. Strossmayer University of Osijek, Croatia

Gregor Rozinaj

Slovak University of Technology, Slovakia

Imre Rudas

Budapest Tech, Hungary

Ivan Samardžić

J.J. Strossmayer University of Osijek, Croatia

Dražan Slišković

J.J. Strossmayer University of Osijek, Croatia

Marinko Stojkov

J.J. Strossmayer University of Osijek, Croatia

Cristina Seceleanu

Mälardalen University, Sweden

Siniša Srblić

University of Zagreb, Croatia

Zdenko Šimić

University of Zagreb, Croatia

Damir Šljivac

J.J. Strossmayer University of Osijek, Croatia

Domen Verber

University of Maribor, Slovenia

Dean Vučinić

Vrije Universiteit Brussel, Belgium

J.J. Strossmayer University of Osijek, Croatia

Joachim Weickert

Saarland University, Germany

Drago Žagar

J.J. Strossmayer University of Osijek, Croatia

Proofreader

Ivanka Ferčec

J.J. Strossmayer University of Osijek, Croatia

Editing and technical assistance

Davor Vrandečić

J.J. Strossmayer University of Osijek, Croatia

Steaphen Ward

J.J. Strossmayer University of Osijek, Croatia

Dražan Bajec

J.J. Strossmayer University of Osijek, Croatia

Journal is referred in:

- Scopus
- Web of Science Core Collection (Emerging Sources Citation Index - ESCI)
- Google Scholar
- CiteFactor
- Genamics
- Hrčak
- Ulrichweb
- Reaxys
- Embase
- Engineering Village

Bibliographic Information

Commenced in 2010.

ISSN: 1847-6996

e-ISSN: 1847-7003

Published: quarterly

Circulation: 300

IJECES online

<https://ijeces.ferit.hr>

Copyright

Authors of the International Journal of Electrical and Computer Engineering Systems must transfer copyright to the publisher in written form.

TABLE OF CONTENTS

Evaluation of objective video quality assessment methods on video sequences with different spatial and temporal activity encoded at different spatial resolutions 1

Original Scientific Paper

Jelena Vlaović | Drago Žagar | Snježana Rimac-Drlje | Mario Vranješ

High-Level Descriptors for Fall Event Detection Supported by a Multi-Stream Network 11

Original Scientific Paper

Sarah Almeida Carneiro | Silvio Jamil Ferzoli Guimarães | Hélio Pedrini

The Effect of Space-filling Curves on the Efficiency of Hand Gesture Recognition Based on sEMG Signals 23

Original Scientific Paper

Panagiotis Tsinganos | Bruno Cornelis | Jan Cornelis | Bart Jansen | Athanassios Skodras

Characterisation of Signal Amplitude-Frequency for Indoor Power Line Communication Channel in the 1 — 30 MHz Broadband Frequencies 33

Original Scientific Paper

Adedayo Oluyomi Ajibade | Ilesanmi Banjo Oluwafemi | Israel Esan Owolabi

Steady-State Analysis Of Single Phase IPM Motors By D-Q Harmonic Balance Method 43

Original Scientific Paper

Asogwa Jude. C. | Obe Emeka. S. | Nnadi Damian. B. | Oti Stephen. E.

Battery energy storage technologies overview 53

Review paper

Zvonimir Šimić | Danijel Topić | Goran Knežević | Denis Pelin

About this Journal

IJECES Copyright Transfer Form

Evaluation of objective video quality assessment methods on video sequences with different spatial and temporal activity encoded at different spatial resolutions

Original Scientific Paper

Jelena Vlaović

J. J. Strossmayer University of Osijek,
Faculty of Electrical Engineering, Computer Science and
Information Technology Osijek,
Kneza Trpimira 2b, Osijek, Croatia
jvlaovic@ferit.hr

Drago Žagar

J. J. Strossmayer University of Osijek,
Faculty of Electrical Engineering, Computer Science and
Information Technology Osijek,
Kneza Trpimira 2b, Osijek, Croatia
dzagar@ferit.hr

Snježana Rimac-Drlje

J. J. Strossmayer University of Osijek,
Faculty of Electrical Engineering, Computer Science and
Information Technology Osijek,
Kneza Trpimira 2b, Osijek, Croatia
rimac@ferit.hr

Mario Vranješ

J. J. Strossmayer University of Osijek,
Faculty of Electrical Engineering, Computer Science and
Information Technology Osijek,
Kneza Trpimira 2b, Osijek, Croatia
mvranjes@ferit.hr

Abstract – With the development of Video on Demand applications due to the availability of high-speed internet access, adaptive streaming algorithms have been developing and improving. The focus is on improving the user's Quality of Experience (QoE) and taking it into account as one of the parameters for the adaptation algorithm. Users often experience changing network conditions, so the goal is to ensure stable video playback with a satisfying QoE level. Although subjective Video Quality Assessment (VQA) methods provide more accurate results regarding user's QoE, objective VQA methods cost less and are less time-consuming. In this article, nine different objective VQA methods are compared on a large set of video sequences with various spatial and temporal activities. VQA methods used in this analysis are: Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), MultiScale Structural Similarity Index (MS-SSIM), Video Quality Metric (VQM), Mean Sum of Differences (DELTA), Mean Sum of Absolute Differences (MSAD), Mean Squared Error (MSE), Netflix Video Multimethod Assessment Fusion (Netflix VMAF) and Visual Signal-to-Noise Ratio (VSNR). The video sequences used for testing purposes were encoded according to H.264/AVC with twelve different target coding bitrates, at three different spatial resolutions (resulting in a total of 190 sequences). In addition to objective quality assessment, subjective quality assessment was performed for these sequences. All results acquired by objective VQA methods have been compared with subjective Mean Opinion Score (MOS) results using Pearson Linear Correlation Coefficient (PLCC). Measurement results obtained on a large set of video sequences with different spatial resolutions show that VQA methods like SSIM and VQM correlate better with MOS results compared to PSNR, SSIM, VSNR, DELTA, MSE, VMAF and MSAD. However, the PLCC results for SSIM and VQM are too low (0.7799 and 0.7734, respectively), for the usage of these methods in streaming services instead of subjective testing. These results suggest that more efficient VQA methods should be developed to be used in streaming testing procedures as well as to support the video segmentation process. Furthermore, when comparing results obtained for different spatial resolutions, it can be concluded that the quality of video sequences encoded at lower spatial resolutions in cases of lower target coding bitrate is higher compared to the quality of video sequences encoded at higher spatial resolutions at the same target coding bitrate, particularly when video sequences with higher spatial and temporal information are used.

Keywords – spatial activity, spatial resolution, temporal activity, video streaming, video quality assessment

1. INTRODUCTION

The major increase in Internet accessibility over the last decade has resulted in high demand of different multimedia content availability that is subject to changing network conditions. The prediction is that In-

ternet video traffic will increase from the current 105 EB per month to 240 EB per month by 2022 and consumer Video on Demand (VoD) traffic will nearly double by 2022 [1]. Various VoD applications and adaptive bitrate (ABR) streaming algorithms have been developed, which has solved some of the problems with adapt-

ing to changing network conditions, but there is still room for improvement. Nowadays ABR algorithms take into account various parameters to adapt to changing network conditions like variations in bandwidth, video segment size, and buffer fullness. Quality of video sequences that are played back to the user should be tested so that the amelioration of ABR algorithms could be verified.

Regardless of the network conditions, users request the highest possible Quality of Experience (QoE), which is still a challenging task. As opposed to QoE, the parameters in the Quality of Service (QoS) specification are selected depending on the type of application and are related to technical aspects. QoS was used to quantify the quality in multimedia services for many years, but the question arose whether the technical parameters of the network correspond to the user's perception of video quality, because QoS does not take into account user's subjectivity. Currently a large amount of video streaming services, including social media, use objective Video Quality Assessment (VQA) methods to measure and control the video quality they deliver to end-users. Objective VQA methods can be also used as inputs to QoE predictors [2-3]. Still the correlation between objective and subjective VQA methods should be more thoroughly analyzed.

Human-Computer Interaction (HCI) researchers were first to point out that QoE takes into account emotions, relationships, context, and expectations [4]. The QoE is defined by ITU-T as the user's subjective acceptance of service [5]. The QoE is also defined as a measure of user's satisfaction or bother with the service [6] which differs from subjective VQA, which is focused on predicting user's responses to visible distortions. Authors in [7] state that QoE is a multidisciplinary area based on engineering and cognitive science, economics, and social psychology. The QoE is affected by various factors that can be divided into human, context and system. Human factors include user preferences, different sensorial and cognitive processes. Context factors are economic and social aspects, as well as time and space in which a service is used. System factors are software (SW) and hardware (HW) limitations of electronic devices that are being used [8]. To ensure the highest possible QoE level, the playback of video content should be seamless, without delays and rebuffering events.

Developers of novel services are taking into account subjective and objective VQA methods that are used to quantify user's QoE [9]. The reliability of an objective VQA method is usually verified and quantified by comparing its results to the results of subjective testing. Objective VQA methods are still a vital part of service testing because subjective testing is time-consuming, expensive and cannot provide the measurement of video quality fast enough [10]. Thus, analysis of objective VQA methods on various video sequences with different spatial and temporal activity and with different spatial and temporal resolution is important to further

optimize ABR algorithms testing procedures. The server side in streaming services stores video sequences in segments encoded with various target coding bitrates and with different spatial resolutions. The client side in streaming services connects the received video segments and prepares the video sequence for playback. Depending on the playback device, spatial resolutions of all video segments have to be adjusted. The idea of this article is to analyze the efficiency of VQA methods on the server side, in order to improve the selection of proper target coding bitrates and spatial resolutions with respect to network conditions, but also the spatial and temporal activity of a given video sequence.

For this article ten different video sequences with different spatial and temporal activity were encoded with various parameters regarding target coding bitrates and spatial resolutions. Nine VQA methods have been tested on encoded video sequences after they were scaled to Full HD spatial resolution. All results were then compared to results acquired from subjective testing.

This article is constructed as follows: related work is given in section 2. Section 3 gives information about test setup, selected coding parameters, target coding bitrates and calculated values of spatial and temporal information of selected video sequences. Section 4 that includes results and discussion is followed by the conclusion.

2. INTRODUCTION

Based on the final video quality score of distorted video sequence which can be determined by a computer or a user, VQA methods can be divided into objective and subjective methods. Objective VQA methods can further be divided into:

- full-reference (FR) objective VQA methods: require full reference video sequence for analysis [10]
- reduced-reference (RR) objective VQA methods: require only a number of features from reference video sequence for evaluation of distorted video sequence quality [11, 12],
- no-reference (NR) objective VQA methods: predict quality without using any reference information and do not need any information about the original video [13, 14].

In this article several FR objective VQA methods are analyzed: Mean Sum of Differences (DELTA), Mean Sum of Absolute Differences (MSAD), Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR) [15], Structural Similarity index (SSIM) [16], MultiScale Structural Similarity Index (MS-SSIM) index [17], Visual Signal-to-noise Ratio (VSNR) [18], Netflix Video Multimethod Assessment Fusion (Netflix VMAF) [15], RR metric Video Quality Metric (VQM) [19].

DELTA uses the mean difference of the color components in the correspondent points of image for quality

assessment. MSAD uses the mean absolute difference of the color components in the correspondent points of image for quality assessment [20].

MSE, DELTA, MSAD and PSNR are VQA metrics that do not take into account Human Visual System (HVS) characteristics. MSE and PSNR are based on statistical processing of the mean value of the square of the pixel difference, thus they do not achieve a high correlation with subjective VQA scores. Metrics that achieve better correlation with subjective VQA scores are SSIM and MS-SSIM because they take into account HVS which is suitable for acquiring structural information. MS-SSIM uses structural distortion calculation but on multiple scales. SSIM and MS-SSIM compare structural features extracted from the video sequences as opposed to comparing the pixel values, [15, 16,18].

VSNR uses low-level properties of HVS of contrast sensitivity and visual masking and mid-level properties of global precedence in the decision-making process. It takes into account the near-threshold and supra-threshold properties of HVS to determine the visual accuracy of the analyzed video. Both the near-threshold and supra-threshold properties are presented as Euclidean distances. VSNR is the linear sum of those distances. The characteristic of VSNR are efficiency regarding the memory requirements and computational complexity [18].

Netflix VMAF is a metric that correlates well with Mean Opinion Score (MOS) due to the fact that it combines multiple elementary video quality features. The estimation model was trained using a large MOS dataset. VMAF focuses on quality degradation resulted from rescaling and compression. The perceived quality score is calculated by computing scores from various VQA algorithms and combining them using a support vector machine [15].

VQM metric uses the Discrete Cosine Transform (DCT) to calculate the distortion in the video sequences. The value of VQM increases with the level of degradation in the analyzed video sequence [19].

Before analyzing the correlation between different objective and subjective VQA methods, it is necessary to investigate what affects the QoE of end-users the most. The QoS used for analyzing various multimedia services depends on the type of service for which the parameters are analyzed. Nowadays it is well known that user's expectation of delivered video quality often do not meet the information acquired from measured network parameters. It is useful to gather information about network parameters to get the knowledge of certain events in the network, but that information will not present the user's experience of those events [21].

Video sequence processing procedure consists of recording or generating the video sequence, coding, compression, transferring, decoding, and reproducing. During those processing procedures different factors can influence both the QoE and the QoS and vary the user's experience. Considering that various technical

and non-technical factors influence the user's experience, optimization and measurement of video quality is an elaborate task [22].

As previously mentioned, there are different technical factors that can decrease the quality of the video sequence in processing procedure. Users generally perceive high contrast video sequences as video sequences with better video quality, whereas video sequences with low brightness, contrast and sharpness as video sequences with low video quality [23]. ITU-T states that subjective testing conducted with a group of at least 15 individuals is the most precise VQA method [24]. For researchers to conduct such testing, ITU-T provides regulation concerning viewing conditions, evaluation procedures, criteria for selecting users and materials and methods of data analysis [25].

Taking into account that subjective VQA methods give more accurate results, it is important to state that the QoE is lower if there is substantial initial delay or there is deterioration in the first temporal segment of the video sequence. Also the QoE is lower in the cases of up-switching between successive video quality levels compared to instantaneous up-switching between quality levels by more than one step. Authors in [26] state that quality level switching among video sequences with different spatial resolutions affects the QoE more than switching among video sequences with different temporal resolutions. Nevertheless, some authors point out the drawbacks of subjective testing like the fact that previous experiences of the participants can compromise the test results [27]. Also, subjective VQA results can be affected by the participants' preferences [28].

Authors in [26] state that results acquired by subjective VQA metric like PSNR do not always correlate with the results acquired from the subjective VQA methods.

Although authors in [18] state that objective VQA metrics like MSE and PSNR have low correlation with subjective testing because they do not take into account the properties of HVS, authors in [29] state that PSNR and SSIM have good correlation with subjective testing.

Authors in [30] compared PSNR and SSIM to MS-SSIM and VMAF. They state that PSNR has the worst results due to the fact that it does not consider perceptual information. SSIM performed somewhat better compared to PSNR especially in cases of images with various supra-threshold distortions. Considering the MS-SSIM has multiscale properties, it performed better than SSIM. VMAF performed the best, but only if Netflix dataset is used because it captures scaling compression artifacts and does not perform well with unseen distortions. Authors in [31, 32] state that MS-SSIM and VQM achieve roughly the same results.

Compared to related work, this article compares results of nine full reference VQA methods to the results of subjective testing using ten video sequences with

different spatial and temporal information encoded on three spatial resolutions and various target coding bitrates. The idea is to use a larger set of video sequences to identify the best VQA method to be used in the future for optimizing video segmentation procedures at server side as well as testing ABR algorithms that use scaling of spatial resolutions.

3. TEST SETUP

Testing conducted for this article was done in order to compare different VQA methods and determine which of them gives the most similar results to those obtained by subjective testing when using video sequences with different spatial and temporal information encoded on different spatial resolutions, scenario often seen in streaming services. Ten various video sequences (often used in testing of ABR algorithms) were selected for testing purposes based on their spatial and temporal activity. All sequences were encoded at three spatial resolutions and various target coding bitrates, from 600 kbps to 12600 kbps. The VQA methods analyzed in this article are MSE, PSNR, VSNR, SSIM, MS-SSIM, DELTA, MSAD, VQM and Netflix VMAF.

Results were acquired using MSU Quality Measurement Tool [20]. All selected video sequences are available in FullHD spatial resolution with 25 fps. The core sequences used in this article are titled: BlueSky (BS), Chimera1102353 (C53), Station2 (S2), PedestrianArea (PA), Chimera1102347 (C47), CosmosLaundromat (CL), ElFuenteDance (ED), MeridianConversation (MC), Skateboarding (SK) and Soccer (SO). BS, S2, PA were obtained from the dataset published in [33]. Other video sequences were obtained from the dataset published in [34].

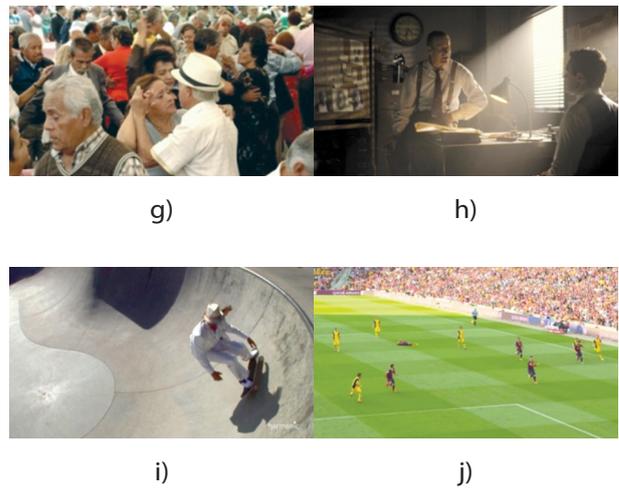


Fig.1. Frames from video sequences: a) Blue Sky (BS) b) Chimera1102353 (C53) c) PedestrianArea (PA) d) Station2 (S2) e) Chimera1102347 (C47) f) CosmosLaundromat (CL) g) ElFuenteDance (ED) h) MeridianConversation (MC) i) Skateboarding (SK) j) Soccer (SO)

Temporal and spatial activity parameters titled Temporal Perceptual Information (TI) and Spatial Perceptual Information (SI) were calculated based on [35] for Y color component of video sequences in YUV format. TI and SI values for all sequences are presented in Fig. 2. From Fig. 2. it can be seen that PA and S2 have similar SI but different TI. Sequences C53 and C47 have similar TI but different SI. The same stands for ED and SO. SO is the sequence with the highest SI, BS is the sequence with the highest TI, whereas C53 is the sequence with lowest SI, MC is the sequence with the lowest TI. This information of SI and TI shall be vital for the analysis of VQA methods.

All selected video sequences were encoded at three spatial resolutions: nHD (640x360), HD (1280x720) and FullHD (1920x1080), according to H.264/AVC video compression standard with coding parameters given in Tab.1, using open-source program called FFmpeg [36]. Spatial resolution downscaling was done using the medium preset and CRF 0 (lossless). FFmpeg uses the scale filter that changes the output sample aspect ratio.

During the coding process, the preset was set to slow because it presents a compromise between time needed for compression and its efficiency. When creating core video sequences, Constant Rate Factor (CRF) was set to zero because it ensures the best possible quality of the output video.

Other sequences encoded from core sequences were encoded using CRF of 23, which is the default value in FFmpeg. Since the CRF was used, the achieved target coding bitrate can vary i.e. be higher or lower than target coding bitrate, because CRF focuses on delivering the requested quality level by using the bitrate close to target coding bitrate. Video sequences with high SI and TI are expected to have the highest achieved coding bitrate for equal target coding bitrate.

All selected target coding bitrates are given in Tab. 2. There are in total 19 different combinations of spatial resolution and target coding bitrate listed in Tab. 2. All ten video sequences with different SI and TI were encoded at those 19 different combinations, thus there are 190 encoded video sequences.

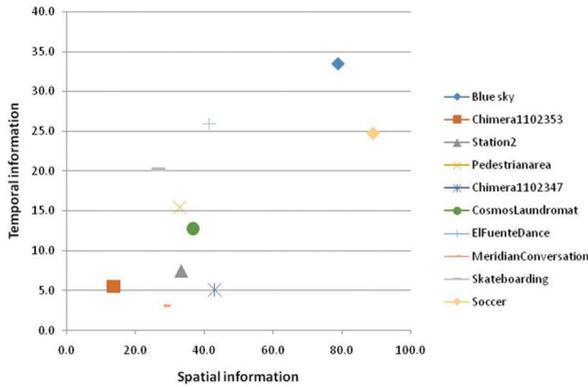


Fig.2. Spatial and temporal activity of video sequences used in our experiments

Tab.1. H.264/AVC coding parameters used in our experiments.

Coding parameter	Value
overall encoding strategy	2-pass variable bitrate
profile	high
coding	CABAC entropy coding
level	4.0
peak rate	1080p Superbit
quantiziercurve compression	0.9
minimum quantization	3

Tab.2. Target coding bitrates.

Spatial resolution	Target coding bitrate [kbps]
nHD	600; 800; 1000; 1400; 2400; 5400
HD	600; 1400; 2400; 3200; 4000; 5600; 9000
Full HD	600; 1400; 5400; 7200; 9000; 12600

In order to compare the quality of sequences with different spatial resolutions, before the objective and subjective quality evaluation, video sequences with nHD and HD resolution were up-scaled to Full HD. Bilinear interpolation was selected as scaling method due to its low complexity but overall satisfying results. The bilinear interpolation was also done using FFmpeg. In this way, we simulated conditions in which ABR may request from the server video segments with a lower resolution due to changes in network throughput, but the video sequence is always presented to a viewer with the same (the highest possible) resolution.

4. RESULTS AND DISCUSSION

Fig. 3. to Fig. 6. present test results for VQA methods PSNR, SSIM, VQM, and Netflix VMAF. All figures show measurement results for ten video sequences encoded at Full HD spatial resolution with 600, 1400, 5400, 7200, 9000, 12600 target coding bitrates. From Fig. 3 it can be seen that video sequences with higher SI and TI (Fig. 2.) have lower values of PSNR due to the fact that they are more complex to encode with the selected coding parameters. Furthermore, it can be concluded that video sequences with higher SI have lower PSNR values in cases when comparing two video sequences with similar TI like C53 and C47. Looking at values from S2 and PA video sequences, it can be concluded that in case of video sequences with similar SI, video sequence with higher TI has lower PSNR values. All video sequences have a value drop at the target coding bit rate of 600 kbps due to exceedingly low target coding bitrate for the analyzed spatial resolution. The lower the values of SI and TI are, the lower the drop in PSNR value is. Similar conclusions considering the relation of measurement results to SI and TI can be gathered from Fig. 4. to Fig. 6. The only main difference is the scale the VQA metric uses, thus the curves look more or less scattered. In addition to objective quality testing, subjective testing was performed for all 190 video sequences. The subjective VQA measurement was done in a controlled environment with 26 inexperienced viewers [37]. After conducting the experiment, MOS was calculated as a mean value of gathered results for each video sequence.

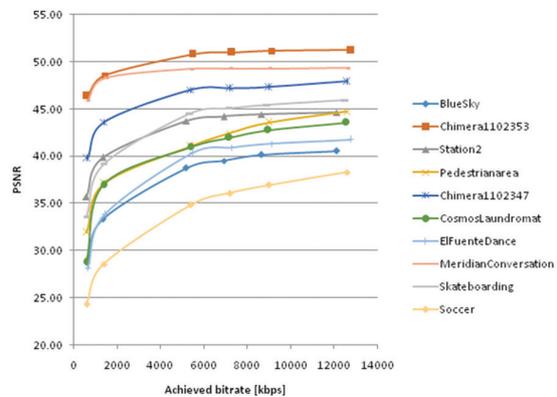


Fig.3. PSNR values for Full HD spatial resolution

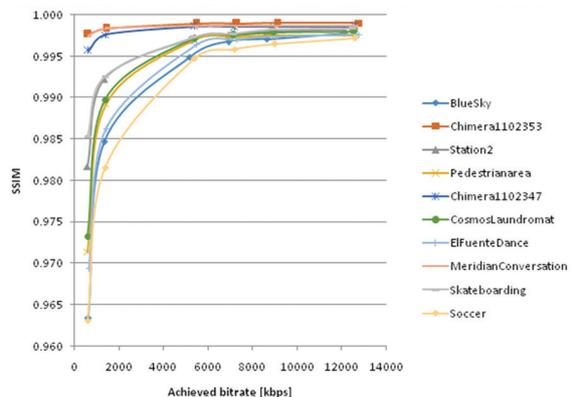


Fig.4. SSIM values for Full HD spatial resolution

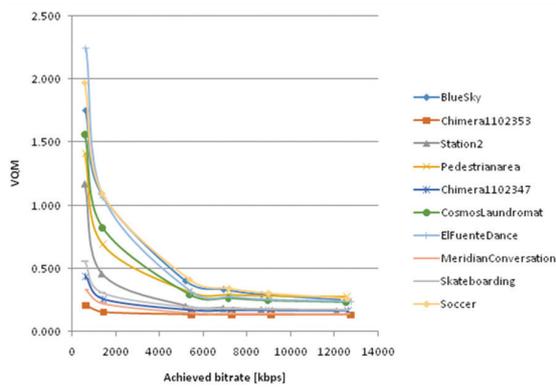


Fig.5. VQM values for Full HD spatial resolution

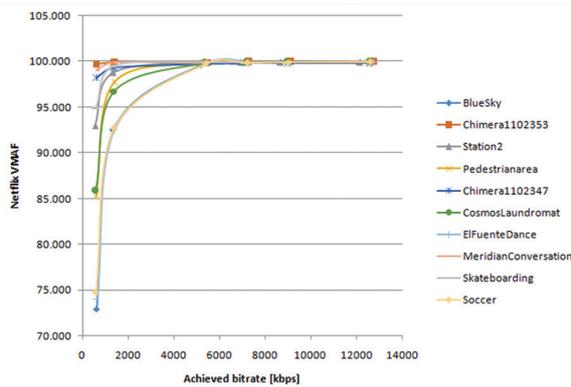


Fig.6. Netflix VMAF values for Full HD spatial resolution

Pearson Linear Coefficient (PLCC) was calculated to determine the correlation between measurement results from all objective VQA methods and subjective test results [38]. PLCC results have been calculated for every spatial resolution for every video sequence separately, then for all video sequences for certain spatial resolution, and for all video sequences for certain objective VQA method.

Results for PLCC for all video sequences at certain spatial resolution and for all video sequences in total for each objective VQA method are given in Tab. 3. PLCC results for nHD and HD have been calculated for the up-scaled video sequences. In total, 190 video sequences have been analyzed for each objective VQA method. The results showed that the correlation between measurement results of subjective and objective VQA is, for most VQA methods, the lowest for video sequences encoded with HD resolution. This can be explained by the fact that the number of target coding bitrates used in the encoding process for this resolution was the largest.

In general, none of the metrics used achieved a very high PLCC value for all video sequences. However, SSIM and VQM have the highest PLCC (0.7794 and 0.7734, respectively), and they achieve overall the best results when different spatial resolutions are used. Although SSIM has the highest PLCC for all sequences, VQM achieves the best results for HD resolution, i.e. when a larger number of target coding bitrates is used.

Considering that SSIM and VQM take into account some of the HVS characteristics, they outperform PSNR, DELTA, MSAD, MSE and VSNR as can be seen from Tab. 3. The unexpectedly low PLCC results are obtained for Netflix VMAF, comparable to PSNR results, which has only a bit lower PLCC. Although Netflix VMAF combines multiple elementary video quality features and has been trained on streaming video sequences, it does not perform well on our experimental setup. Partially, this can be explained with different settings for subjective measurements used for VMAF development, i.e. model for VMAF is based on the assumption that the video sequences are presented in 1080p resolution TV display and that viewers are on the viewing distance of 3x the screen height (3H).

Tab.3. PLCC results for video sequences at nHD, HD and Full HD spatial resolutions.

Objective metric	Spatial resolution	PLCC
PSNR	360p	0.685044
	720p	0.644172
	1080p	0.715153
	All video sequences	0.708182
SSIM	360p	0.888921
	720p	0.715210
	1080p	0.831389
	All video sequences	0.779926
MS-SSIM	360p	0.718876
	720p	0.681271
	1080p	0.769829
	All video sequences	0.711381
VQM	360p	0.867630
	720p	0.755780
	1080p	0.831150
	All video sequences	0.773982
DELTA	360p	0.823293
	720p	0.634594
	1080p	0.830959
	All video sequences	0.649855
MSAD	360p	0.608980
	720p	0.605221
	1080p	0.684087
	All video sequences	0.699860
MSE	360p	0.726784
	720p	0.691223
	1080p	0.697680
	All video sequences	0.666198
Netflix VMAF	360p	0.670822
	720p	0.585973
	1080p	0.718319
	All video sequences	0.718851
VSNR	360p	0.543860
	720p	0.525684
	1080p	0.712838
	All video sequences	0.708905

In our experiment, viewers watched video sequences on a computer monitor at a distance of less than 3H. When analyzing the results for each video sequence it can be concluded that objective VQA method results obtained for video sequences with higher spatial and temporal information have lower PLCC which states for all VQA methods (Tab. 4. presents SSIM results for C53, CL and SO, but similar results were obtained for all VQA methods).

The lower PLCC values for video sequences with high SITI (a product of SI and TI values, $SITI=SI \cdot TI$), are a result of spatial and temporal masking of the errors that occur due to a large number of details and fast motions in video sequences with high SI and TI. Objective methods compare video sequences frame by frame (full reference metrics) and do not take properly the effects of spatial and temporal masking into account, so they generally rate the quality more rigorously compared to human viewers for these fast and complex sequences. That can cause the metrics to overestimate the visible impairments and give lower objective VQA scores compared to subjective scores.

Taking into account that ABR algorithms use video segments with different spatial resolution, analysis has been done on results given for every spatial resolution separately. Fig.7. presents SSIM results for video sequences CL, C53 and SO for all three spatial resolutions. CL, C53 and SO are selected because C53 has the lowest SITI, SO has very high SITI and the highest SI, and SITI of CL is in the middle compared to C53 and SO (Fig. 2.). Although results are given for SSIM, all other metrics give similar results. From Fig. 7. it can be concluded that in general the SSIM values are higher when achieved bit rate is higher and that video sequences encoded at higher spatial resolution have higher SSIM values. Still, in cases when spatial and temporal activity of video sequence are too high to encode it on low target coding bitrate and with HD or Full HD spatial resolutions, SSIM values can be higher in the case of nHD spatial resolution. For example, in the case of the SO video sequence that has high spatial and temporal activity, SSIM values for the target bit rate of 600 kbps are lower for both HD and Full HD compared to nHD. In the case of CL that has a lower SI and TI at 600 kbps, SSIM value for HD spatial resolutions is higher than for Full HD spatial resolution. Taking into account SI and TI, it can be seen that the bit rate at which the overlap occurs is higher for video sequences with higher SI and TI. This conclusion can help with selecting the most suitable target coding bitrates for each spatial resolution, thus improving encoding and segmentation process when preparing video sequences for adaptive streaming. For the C53 sequence, which has the lowest SITI, there are no overlapping effects for nHD resolution. SSIM results for this resolution are lower than for HD and Full HD even for 600 kbps, because due low spatial and temporal activity of this video sequence, coder can successfully encode at higher spatial resolutions. It can be expected that overlapping for sequences with such low SITI occurs at even lower target coding

bitrates. The MOS results for video sequences C53, CL and SO, given in Fig. 8, confirm that the overlapping of curves occurs. For CL and SO video sequences it can be seen that at lower target coding bitrates, MOS as well as SSIM can be higher for lower spatial resolution. The C53 sequence has much lower MOS results at nHD than at HD and Full HD resolutions, even at 600 kbps, because the loss of details caused by a decrease in spatial resolution cannot be masked due to the low spatial and temporal activity of the video content.

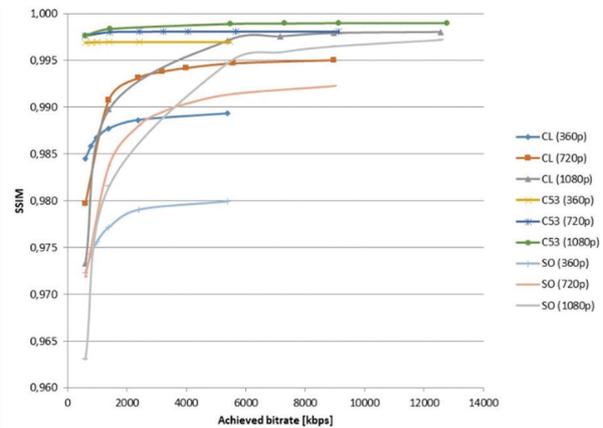


Fig.7. SSIM values for three spatial resolutions for video sequences Chimera1102353 (C53), CosmosLaundromat (CL) and Soccer (SO)

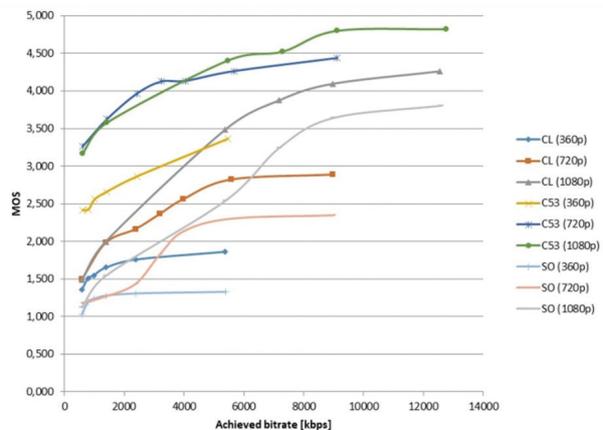


Fig.8. MOS values for three spatial resolutions for video sequences Chimera1102353 (C53), CosmosLaundromat (CL) and Soccer (SO)

Tab. 4. PLCC results for SSIM and SITI for Chimera1102353 (C53), CosmosLaundromat (CL) and Soccer (SO) video sequences.

Objective metric	Video sequence	SITI	PLCC
SSIM	Chimera1102353 (C53)	75.066	0.921
	CosmosLaundromat(CL)	470.342	0.842
	Soccer (SO)	2200.608	0.821

5. CONCLUSION

In order to improve datasets and the segmentation process for video streaming purposes in the future, different variants of encoded video sequences were used. In this paper, nine subjective VQA methods were analyzed using ten video sequences with different SI and TI that were encoded at three spatial resolutions and various target coding bitrates. From our results, video sequences with higher SI and TI have lower values of PSNR. When comparing the sequences with similar TI, it can be concluded that sequences with higher SI have lower PSNR values. Sequences with higher TI and similar SI have lower PSNR values. PSNR values drop considerably at coding bit rate of 600 kbps, especially for video sequences with higher SI and TI. Similar results are obtained for all analyzed objective VQA metrics. Results also show that in cases of higher SI and TI and low target coding bitrate, subjective VQA scores can be higher for lower spatial resolutions. The same conclusion can be made from MOS results. On selected dataset, SSIM achieves the best overall correlation to PLCC results calculated based on MOS thus it is the best in cases when video sequences are encoded with various spatial resolutions and various target coding bitrates. Video sequences with higher SI and TI have lower PLCC results due to spatial and temporal masking, which objective VQMs fail to capture well. VQM acquires the best results for HD resolution which was calculated for larger set of encoded sequences, though SSIM has the highest overall PLCC. Compared to PSNR, DELTA, MSAD, MSE and VSNR, SSIM and VQM have higher values of PLCC because they consider HVS characteristics. From MOS results and results obtained with objective VQA methods it can also be concluded that when sequences with high SI and TI encoded at low target coding bitrate are used, video quality can be higher in case of nHD spatial resolution compared to HD and Full HD spatial resolutions. This situation does not occur when video sequences with low SI and TI are used. In the future work those conclusions shall be used to shape parameters for new adaptive streaming algorithm and to propose a new dataset for adaptive streaming purposes.

6. ACKNOWLEDGMENT

The current archival periodical article is based on the conference presentation [37].

7. REFERENCES

- [1] Cisco Visual Networking Index: Forecast and Methodology, 2016–2021, San Jose, CA, USA, Sep. 2017.
- [2] C. G. Bampis and A. C. Bovik. (2018). "Feature-based prediction of streaming video QoE: Distortions rebuffering and memory." *Signal Processing: Image Communication*. vol. 68, no. 7, pp. 218–228, Oct. 2018.
- [3] C. G. Bampis, Z. Li, I. Katsavounidis, and A. C. Bovik, "Recurrent and dynamic models for predicting streaming video quality of experience," *IEEE Trans. Image Process.*, vol. 27, no. 7, pp. 3316–3331, Jul. 2018.
- [4] Y. Chen, K. Wu, and Q. Zhang, "From QoS to QoE: A tutorial on video quality assessment," *IEEE Communications Surveys & Tutorials*, vol. 17, pp. 1126–1165, October 2015.
- [5] "ITU-T G.1080 Quality of experience requirements for IPTV services," *International Telecommunication Union*, pp. 1-44, 2008.
- [6] P. L. Callet, S. Moller, and A. Perkis, "Qualinet white paper on definitions of quality of experience (2012)," *Eur. Netw. Qual. Exp. Multimedia Syst. Services (COST Action IC 1003)*, White Paper, 2013.
- [7] G. Pibiri, C. Mc Goldrick, and M. Huggard, "Expected Quality of Service (eQoS) A network metric for capturing end-user experience," *IFIP Wirel. Days*, November 2012.
- [8] U. Reiter et al. "Factors influencing Quality of Experience," in *Quality of Experience*, S. Moeller and A. Raake, Eds. London: Springer, 2014, pp. 55-72.
- [9] H-J. Park and D-H. Har, "Subjective Image Quality Assessment based on Objective Image Quality Measurement Factors," *IEEE Trans. On Consumer Electron.*, vol. 57, no. 3, pp. 1176-1184, Aug. 2011.
- [10] S. Chikkerur, V. Sundaram, M. Reisslein, and L. J. Karam, "Objective video quality assessment methods: A classification, review, and performance comparison," *IEEE Trans. Broadcast.*, vol. 57, no. 2, pp. 165–182, Jun. 2011.
- [11] Z. Wang, L. Lu, and A. C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal Process., Image Commun.*, vol. 19, no. 2, pp. 121–132, 2004.
- [12] Z. Wang and E. P. Simoncelli, "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model," *Electron. Imag.*, vol. 5666, pp. 149–159, Mar. 2005.
- [13] A. Mittal, M. A. Saad, and A. C. Bovik, "A completely blind video integrity oracle," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 289–300, Jan. 2016.
- [14] Nighi, N. Aggarwal, "A review on Video Quality Assessment", 2014 Recent Advances in Engineering and Computational Sciences (RAECS), Chandigarh, 2014, pp. 1-6, March 2014.
- [15] R. Rassool, "VMAF reproducibility: Validating a percep-

- tual practical video quality metric," 2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), pp. 1-2, June 2017.
- [16] Z. Wang, A. Bovik, and H. Sheikh, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, No. 4, Apr 2004.
- [17] C. Yang, L. Zhao and Z. Liao, "Objective Quality Metric Based on Perception for Video," *International Conference on Computer Engineering and Technology*, pp. 20-23, January 2009.
- [18] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Transactions on Image Processing*, vol. 16, pp. 2284-2298, August 2007.
- [19] M. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *IEEE Trans. Broadcast.*, vol. 50, no. 3, pp. 312-322, Sep. 2004.
- [20] D. Vatolin, M. Smirnov, A. Ratushnyak and V. Yockin, www.compression.ru/video/quality_measure/video_measurement_tool.html (accessed: 2020)
- [21] P. Orosz, T. Skopkó, Z. Nagy, P. Varga, and L. Gyimóthi, "A case study on correlating video QoS and QoE," *IEEE Network Operations and Management Symposium (NOMS)*, pp. 1-5, May 2014.
- [22] V. Deksnys, E. Sakalauskas and G. Činčikas, "New integral quality of TV service criterion construction based on quality of experience statistical estimation," *14th Biennial Baltic Electronic Conference (BEC)*, pp. 149-152, October 2014.
- [23] R. Koshimura, Y. Ito and Y. Nomura, "Empirical study on clarification of relationship between QoS and QoE for Web services by path analysis," *3rd Global Conference on Consumer Electronics (GCCE)*, pp. 10-11, 2014, October 2014.
- [24] "ITU-T P.10 Vocabulary for performance, quality of service and quality of experience," *International Telecommunication Union*, pp. 1-22, 2017.
- [25] "ITU-T P.912: Subjective video quality assessment methods for recognition tasks," *International Telecommunication Union*, pp. 1-22, 2016.
- [26] T. Zinner, O. Hohlfeld, Osama Abboud and T. Hossfeld, "Impact of Frame Rate and Resolution on Objective QoE Metrics," *Second International Workshop on Quality of Multimedia Experience (QoMEX)*, pp. 29-34, June 2010.
- [27] P. Orosz, T. Skopkó, Z. Nagy, P. Varga and L. Gyimóthi, "A Case Study on Correlating Video QoS and QoE," *IEEE Network Operations and Management Symposium (NOMS)*, pp 1-5, May 2014.
- [28] D. Z. Rodríguez, R. L. Rosa, E. A. Costa, J. Abrahão, and G. Bressan, "Video Quality Assessment in Video Streaming Services Considering User Preference for Video Content," *IEEE Transactions on Consumer Electronics*, vol. 60, pp. 570 - 571, March 2014.
- [29] M. Vranješ, S. Rimac-Drlje, D. Žagar, "Subjective and Objective Quality Evaluation of the H.264/AVC Coded Video," *15th International Conference on Systems, Signals and Image Processing*, pp. 287 - 290, June 2008.
- [30] C. G. Bampis, Z. Li and A. C. Bovik, "Spatiotemporal Feature Integration and Model Fusion for Full Reference Video Quality Assessment," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 8, pp. 2256-2270, Aug. 2019.
- [31] D. Vranješ, D. Žagar and O. Nemčić, "Comparison of objective quality assessment methods for scalable video coding," *Proceedings ELMAR-2012*, pp. 19-22, September 2012.
- [32] Z. Duanmu, K. Ma and Z. Wang, "Quality-of-Experience for Adaptive Streaming Videos: An Expectation Confirmation Theory Motivated Approach," *IEEE Transactions on Image Processing*, vol. 27, no. 12, pp. 6135-6146, 2018.
- [33] <https://media.xiph.org/> (accessed: 2020)
- [34] C. G. Bampis, Z. Li, I. Katsavounidis, TY Huang, C. Ekanadham and A. C. Bovik, "Towards Perceptually Optimized End-to-end Adaptive Video Streaming," submitted to *IEEE Transactions on Image Processing*.
- [35] "ITU-T P.911 Subjective audiovisual quality assessment methods for multimedia applications," *International Telecommunication Union*, pp. 1-27, 1998.
- [36] <https://www.ffmpeg.org/> (accessed: 2020)
- [37] J. Vlaović, M. Vranješ, D. Grabić and D. Samardžija, "Comparison of Objective Video Quality Assessment Methods on Videos With Different Spatial Resolutions," *2019 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pp. 287-292, June 2019.
- [38] W. Kirch, *Pearson's Correlation Coefficient*, Springer, 2008.

High-Level Descriptors for Fall Event Detection Supported by a Multi-Stream Network

Original Scientific Paper

Sarah Almeida Carneiro

Institute of Computing, University of Campinas
Campinas, SP, 13083-852, Brazil
sarah.alcar@gmail.com

Silvio Jamil Ferzoli Guimarães

Computer Science Department, Pontifical Catholic University of Minas Gerais (PUC Minas)
Belo Horizonte, MG, 30535-065, Brazil
sjamil@pucminas.br

Hélio Pedrini

Institute of Computing, University of Campinas
Campinas, SP, 13083-852, Brazil
helio@ic.unicamp.br

Abstract – *The need for assertive video classification has been increasingly in demand. Especially for detecting endangering situations, it is crucial to have a quick response to avoid triggering more serious problems. During this work, we target video classification concerning falls. Our study focuses on the use of high-level descriptors able to correctly characterize the event. These descriptor results will serve as inputs to a multi-stream architecture of VGG-16 networks. Therefore, our proposal is based on the analysis of the best combination of high-level extracted features for the binary classification of videos. This approach was tested on three known datasets, and has proven to yield similar results as other more consuming methods found in the literature.*

Keywords – *Video Classification, Multi-stream Network, Fall Detection, High Level Features, Convolutional Neural Network*

1. INTRODUCTION

Abnormal situations can be characterized as actions that fall outside the scope of a given context. Therefore, an example of such, can be the identification of endangering situations. A condition that can be considered a risk, especially for older people, are falls. Due to the weakening of the body's physical structures, an elderly person's fall can lead to various other problems that can contribute to more severe outcomes. Accordingly, it is critical the fast and correct identification of these actions to avoid further complications.

A way of determining some of these abnormal actions is through the analysis of videos from surveillance cameras. Thus, the identification of these cases in videos is being facilitated by the use of computational resources. The sole employment of a surveillance camera operator is not always as assertive as an on-going surveillance algorithm. This can be observed since people can get easily distracted, unlike a machine.

The use of neural network approaches has been playing an important role in action identification in videos.

There are many branches considering these studies, one of them is video classification. Video classification is the verification of the existence of a given action, in a group of actions, in an analyzed input video.

This study, an extension of Carneiro et al. [1], still influenced by the identification of falls, focuses on the binary classification of videos. Since many works use information that is directly linked to the RGB information, such as the video frame itself, our proposal is based on the use of data only generated from a high-level descriptor extraction. It is a concern of ours that the RGB can influence the classification behavior of the network when dealing with the generalization of cases.

The use of a specific dataset, that is influenced by camera noise, actors, furniture and others, can make the network not as assertive as expected if the frame information is solely used. In addition, since high-level descriptors might not be as assertive as an RGB information for a given dataset, we also focus on the merge of three of these descriptors: (i) optical flow; (ii) visual rhythm; and (iii) pose estimation. Although some of these features have already been used in other works, their combina-

tion is a novel approach. This combination can provide us with complementary temporal, spatial and rhythmic (temporal-spatial) information of the video without having to rely on the raw frame information itself.

This descriptor ensemble was thought of as a three-stream VGG-16 [2] architecture known as a multi-stream. With this multi-stream, we were able to verify the best complementary stream combination for a video classification concerning falls. Furthermore, by avoiding the use of RGB, we are able to conceal the identity of the people in the analyzed videos as well as to observe that the combination of lesser information descriptors can provide as good results as the features commonly used in the literature.

This text is organized as follows. In Section 2, we discuss some of the recent works associated with fall detection in videos. In Section 3, concepts used in this work are clarified. In Section 4, the proposed methodology is explained. In Section 5, we describe the experiments performed and compare the achieved results to other published methods. Finally, some concluding notes and suggestions for future work are presented in Section 6.

2. RELATED WORKS

Given the aging process, it is possible to notice a weakening of various body structures. Associated with this, it is observable that as well as reflexes, balance can also be affected. Therefore, these issues allied to a number of other factors can be responsible for the occurrence of falls. In addition, since recovery for this portion of the population might not be as fast, falling situations might lead to aggravated injuries. Thus, it is imperative to have access as quickly as possible to aid. Hence, several researches associated with computational resources, related to wearable sensors, video processing and machine learning, have been conducted to try to identify these falls faster.

A study associated with video information was conducted by Lin et al. [3] regarding the extraction of motion vectors and DC+2AC images. This information is used as input to a Global Motion Estimator (GME) to cluster global and local motion. Based on this clustering, falls can be identified by the analysis of the person's centroid, the vertical projection histogram value, and the event duration. There is also research related to identifying falls based on silhouette recognition. These studies use this data as input to statistical models, such as Hidden Markov Model (HMMs), to determine the occurrence of falls [4][5].

It has also become important for these studies to conceal the subject's identity. Accordingly, privacy can be achieved by blurring, silhouetting, covering the object with graphical shapes, among other strategies. A common element that is also associated with action detection is background subtraction. Thus, a subtraction method in conjunction to head tracking algorithms was proposed by Yu et al. [6] to identify falls. As observed, head tracking is also useful in this scenario,

Yu et al. [7] was able to correlate this data to density calculation methods, with a mixture Gaussian model for fall detection. Similarly, with a Gaussian model, Rougier et al. [8] was able to use a feature of human body deformity to cope with falling. Attempts for detection were also made by using a subject's bounding box surrounding angulations to train a K-Nearest Neighbor (KNN) [9].

Depth information also started to be considered among detection studies. The RGB-D (RGB image and analogous depth image) were used to extract features, such as Histogram of Oriented Gradients (HOG), Optical Flow (OF) and target skeletons. These features were served to a Support Vector Machine (SVM) to be able to classify fall events [10]. It has also been a concern to determine whether a learning algorithm has a good performance associated with action detection. Thus, a comparative research was conducted to verify distinct learning structures associating them to falls [11].

Kwolek and Kepski [12] demonstrated that, in given circumstances, the KNN algorithm could provide better results compared to the SVM algorithm. The optical flow is regularly associated with a VGG for classification scenarios. By using it, Nunez-Marcos et al. [13] classified falls with a three-stage transfer learning method. Naive-Bayes has also been used as an alternative for detection, employing as data Bag-of-Words from silhouette oriented volumes (SOV) strategy [14]. Alternative studies used a modified CNN AlexNet architecture allied to transfer learning applied to fall detection for surveillance videos [15]. The use of curvelet transforms and an SVM for the identification of human postures are also used with a hidden Markov model (HMM) to classify the existence of falls in videos [16].

3. THEORETICAL BACKGROUND

In this section, we will clarify some of the basic knowledge necessary for understanding this work.

3.1. Optical Flow

The optical flow is a high level feature descriptor able to outline an approximation of a possible movement between video frames (Figure 1). This identified movement can, for example, be caused by an object displacement or a camera shift. Therefore, the algorithm's objective is to highlight frame information that was relocated from a previous frame to its following. Hence, we are able to obtain only the moving object information in the image, thus, being able to discard every other extra static information in the frame.

Let $l(x,y,t)$ be a pixel in a video's first frame $f1$ and d_t the elapsed time between the first frame and its next compared frame $f2$. Considering that the pixel $l(x,y,t)$ was relocated by a distance (d_x, d_y) , based on the condition that the pixels of $f1$ and $f2$ have equivalent static intensities, we are able to use Equation 1, where x, y reference the coordinate of a pixel at the t positioned frame. Then, after applying a Taylor series approxima-

tion, it is possible to calculate the optical flow with Equation 2 through 6. There are two possible ways to calculate the optical flow of a set of frames, by a sparse or a dense methodology.

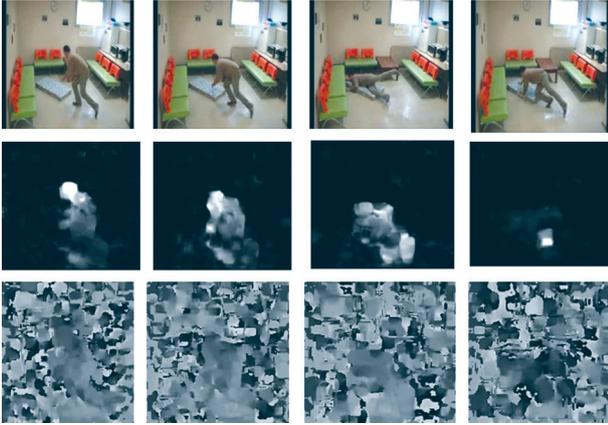


Fig. 1. Frame examples from the FDD dataset and their optical flow extraction example in y and x components, respectively.

A sparse application will generate the optical flow information for selected pixels that can be considered relevant for an object, for instance edges. The dense method, on the other hand, calculates the optical flow per pixel. Despite being a more expensive approach, our study used the Farneback [17] algorithm to generate the video's optical flow filtered frames, which is a dense optical flow calculation. This means that the flow vectors are calculated for each pixel. However, this choice was made once we

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad (1)$$

$$I(x + dx, y + dy, t + dt) \approx I(x, y, t) + \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt \quad (2)$$

$$f_x = \frac{\partial I}{\partial x} \quad f_y = \frac{\partial I}{\partial y} \quad f_t = \frac{\partial I}{\partial t} \quad (3)$$

$$f_x dx + f_y dy + f_t dt = 0 \quad (4)$$

$$u = \frac{dx}{dt} \quad v = \frac{dy}{dt} \quad (5)$$

$$f_x u + f_y v + f_t = 0 \quad (6)$$

3.2. Visual Rhythm

The visual rhythm of a frame-set is a descriptor capable of providing spatio-temporal information about the video [18-20]. Although not visually intuitive, this descriptor, by gathering information of each individual video frame, is able to concatenate this information to generate an outcome that is composed of a single image capable of summarizing the video used as input.

This descriptor can be built based on a variety of methods. Therefore, different outcomes can be generated depending on the approach used.

A first approach can be by using histograms. Consider a video $V=(f_t)_{t \in [0, T-1]}$ in the $2D + t$ domain, $D = \{0, \dots, M-1\} \times \{0, \dots, N-1\}$, where M and N are the width and height of the

frame. The variable f_t is a progression of frames and T is the total number of frames in V . A visual rhythm ϑ can be produced, for example, by assembling a sequence of histograms $(H_{f_t})_{t \in [0, T-1]}$ calculated based on all the frames of a video. ϑ is defined in Equation 7, where $Z \in [0, L-1]$ and $t \in [0, T-1]$, such that T is the number of frames and L the number of histogram bins. Therefore, the result is a 2D representation of the combination of all frame histograms, where each column of ϑ represents a frame histogram. This process is exemplified in Figure 3.

$$\vartheta(t, z) = H_{f_t}(z) \quad (7)$$

Another technique that can be used to generate a visual rhythm is by sub-sampling [21][22]. Hence, the visual rhythm, in domain $1D+T$, is a rendition of the video $V=(f_t)_{t \in [0, T-1]}$ in which each frame f_t has their pixels sampled into a column of ϑ (Equation 8). Accordingly, $Z \in [0, \dots, h_{\vartheta}-1]$ and $k \in [0, \dots, w_{\vartheta}-1]$, where z , k , r_x and r_y are the height and the width of ϑ , and the ratios of pixel sampled from f_t respectively. In addition, A and B relate to the shifts on each frame, from where the algorithm can initiate the pixel sampling (Figure 3).

$$\vartheta(k, z) = f_t(r_x \times z + a, r_y \times z + b) \quad (8)$$

Although the output of these algorithms is a single image for a video, in the attempt of gathering more spatio-temporal information, we modified the algorithm so that it could produce a visual rhythm for each frame. Thus, we are able to define how many frames we intend to use to create a visual rhythm that is associated with f_t . Consequently, considering a window of 5 frames, we use f_t plus its 5 following frames to calculate a visual rhythm and associate it with f_t , and so on until the last frame.

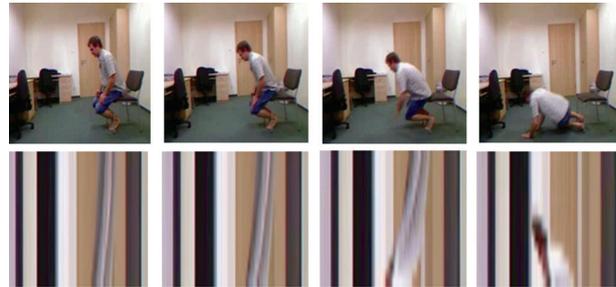


Fig. 2. First row: Frame examples from the URFD dataset Second row: The frame visual rhythm extraction examples.

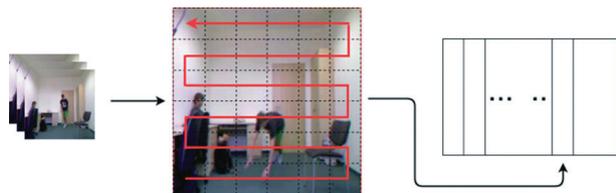


Fig. 3. Visual rhythm algorithm example.

The red arrow is demonstrating a possible way of pixel sampling of a frame. The information gathered is placed as a column of the visual rhythm output image.

3.3. Pose Estimation

The pose estimator used during this work is a descriptor able to detect a humanoid and determine its main joints. Similar to the optical flow, the pose estimation will conceal unnecessary information for a classification study (Figure 4).

For this descriptor, we used the algorithm proposed by Cao et al. [23]. The approach receives a video frame that is then used on a two-branch CNN feed-forward network, that is, networks that do not return the information calculated by a layer to a previous one. Thus the information runs through the network until it reaches and becomes an output. An example is a multi-layer perceptron, that is able to predict sets of 2D confidence maps (S) of body member positioning (where might a body member be positioned in the frame) and 2D vector fields (L) (that calculates the degree of association between these members).

The set $S = (S_1, S_2, \dots, S_J)$, where J is the number of body parts found, and $L = (L_1, L_2, \dots, L_C)$, where C is the number of limbs, are parsed by bipartite matching greedy inference so they can be associated. Lastly, the algorithm produces the 2D key-points indicating the posture of all human components of the frame.

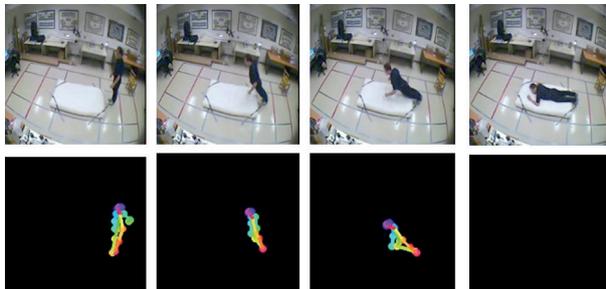


Fig. 4. Frame examples from the Multicam dataset and their pose estimation examples.

3.4. Multi-Stream Architecture

The multi-stream architecture [28-30] is a learning algorithm based on the ensemble of individual learners, each learner is considered a stream (Figure 5). The ensemble can be made based on some different approaches, such as an average of the individual results, the use of a support vector machine (SVM), and others [24]. The use of a multi-stream makes it possible for the network to learn each of the high-level descriptors without disregarding a descriptor as unimportant. Another feature is that, with multi-stream, we are able to evaluate the use of all descriptor combinations. In other words, considering the employment of three high-level descriptors, we can analyze the results of a combination one by one, two by two, and all three.

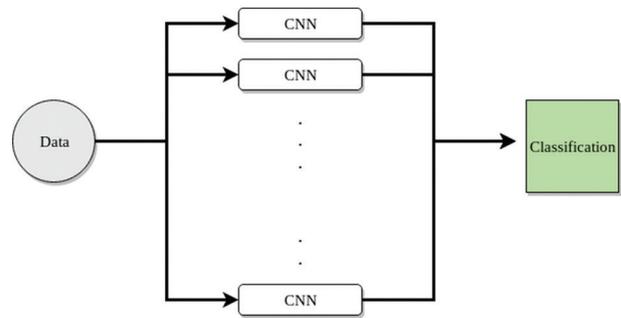


Fig. 5. Multi-stream architecture example.

4. MULTI-STREAM FALL DETECTION APPROACH

Since previous studies regarding abnormal action detection in videos have used standard descriptors for classification, our goal was to verify if the use of weaker feature descriptors could provide as satisfying classification results concerning fall detection. The purpose of this is based on the fact that, by discovering good descriptors, we are able to manage potentially faster or, even, more specialized descriptors for classifying each abnormal situation depending on its need.

The methodology proposed in this work is based on an ensemble of three VGG-16 streams. We acknowledge that the use of the VGG-16 might be outdated, however, it is a classical method and have yielded satisfying results regarding classification in past studies. In addition, since the goal was to verify the relevance of feature combination rather than faster learning algorithms we decided to keep this architecture for a primary result analysis. The features chosen for this study were the optical flow, the visual rhythm, and the subject's pose estimation. Even though some of these features have already been used in past literature works, the proposition of their combination is, to our knowledge, novel.

4.1. High-Level Descriptor Extraction

In this work, we focused on three high-level descriptors: (i) Optical Flow; (ii) Pose Estimation; (iii) Visual Rhythm. As an initial step of our method, the videos from each dataset needed to go through the investigated descriptor algorithms. This pre-processing step guarantees the original frame information filtering according to our need of temporal, spatial or spatial-temporal data. From each frame, these extractions generate other 2D image information (Figures 1, 2 and 4). These images will then serve as inputs to what we acknowledge as a high-level feature extraction.

4.2. High-Level Feature Extraction

After obtaining the images provided by the descriptors, each of these image groups (Optical Flow, Pose Estimation, Visual Rhythm) go through a modified VGG as it can be observed in Figure 6. It should be noticed that

this modified CNN is not a part of the stream learning structure. It is mainly used to avoid explicit feature engineering, since it is not intuitive or easy to determine which are the best features to extract and to ensure the method's independence. The extracted feature vectors acquired by the modified VGG serve as weights for the following step, fine-tuning.

4.3. Individual Streams

Our work was based on the feature extraction of three descriptors. Therefore, the learning model needed to have three individual streams. Each of these streams go through a three step learning process: (i) multi-stream pre-training; (ii) fine-tuning; and (iii) ensembling.

In this work, the falling data provided by the datasets did not optimally contribute with the information required by our network. Therefore, more data was needed to train this learning model. Thus, the initial step was to train the first 14 layers of each of the VGG-16 on ImageNet [25] and, later, the UCF101 dataset [26]. This makes it possible for the network to understand the image's basic structures as well as to identify movement and sequences.

To guarantee that the network would learn the falling action itself, as a second step, we then froze these pre-trained layers and fine-tuned the remaining two. This fine-tuning procedure was associated with the feature vectors extracted (in Section 4.2) from each of the previously presented descriptors. Hence, since we implemented three streams, each individual stream was linked to a specific descriptor. In other words, each stream was fine-tuned with only one of the calculated feature vectors. The third and final step is the ensemble of the results of the individual classifications to combine their isolated outcomes regarding the best combination of features for fall classification.

4.4. Classification

To provide the final classification verdict, as mentioned in the previous section, we used ensembling techniques. To analyze which type of ensemble would provide the best results for our classification, we used three different techniques: (i) average and threshold; (ii) average and a support vector machine (SVM); and (iii) continuous values and SVM.

The first technique, average and threshold, computed the average based on the sum of each stream output and compared it to an empirically defined class threshold. The second ensembling method, on the other hand, had this class threshold defined and adjusted based on the assistance of an SVM. Finally, the continuous values and SVM classified the input regarding a generated vector with each of the streams output so that an SVM could find its separation region.

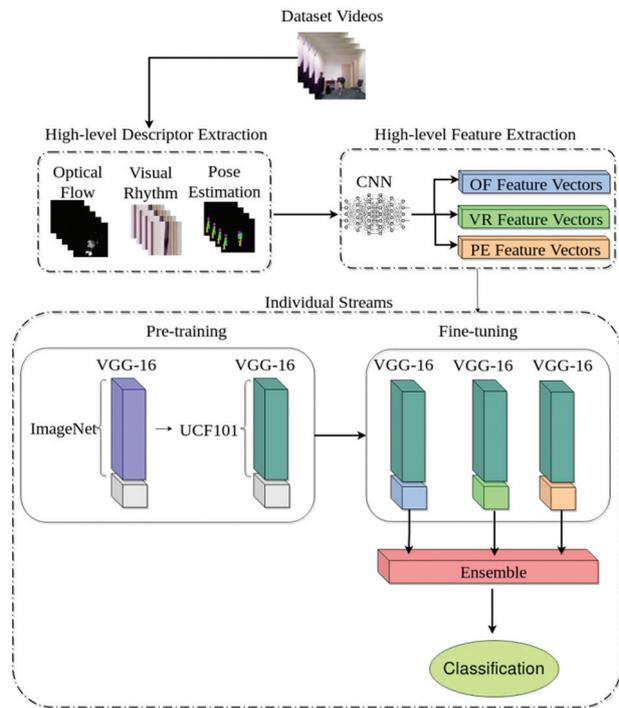


Fig. 6. Multi-stream architecture for fall detection

To analyze the impact of features from high-level descriptors regarding falls, all possible combinations of streams were explored in this work. Therefore, we organized the stream ensembling considering single, two by two, and all three stream results (Figure 7).

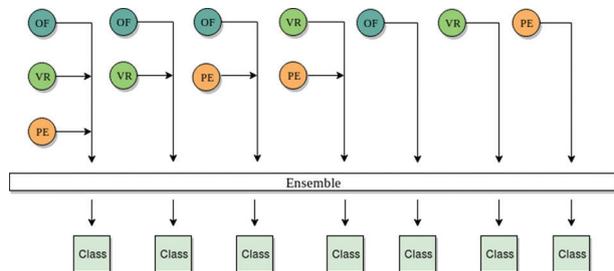


Fig. 7. Feature stream combination.

5. EXPERIMENTS

In this section, we briefly describe the datasets used in our experiments, as well as the quantitative and qualitative results.

5.1. Dataset

Our method was evaluated on three well known falling sets throughout the literature: FDD, URFD [27] and Multicam datasets. These datasets are focused on binary falling and not falling classes. The URFD dataset is composed of 70 videos: (i) 30 videos of falls; and (ii) 40 videos displaying diverse activities. FDD disposes of 191 single camera surveillance videos from both elderly home environments, and office rooms. The Multicam contains 24 scenarios, in which 22 have falling situations, recorded with 8 differently positioned video cameras around the room.

5.2. Quantitative Results

Our quantitative analysis was based on sensitivity, specificity and accuracy metrics. They were chosen since previous studies validated their works using them and we can better evaluate the outputs of our learner. Accuracy is a metric for the model performance evaluation that correlates both positive and negative classes and measures how accurate are the learning results (Equation 9). In Equations 9 to 11, variables TP , TN , FP , FN , P , N stand for true positives, true negatives, false positives, false negatives, total number of positive, and total of negative samples, respectively.

$$A=(TP+TN)/(P+N) \quad (9)$$

Specificity is a metric that provides information related to, given a negative example, the probability of a result being negative (Equation 10). Sensitivity, also known as recall, is, given a positive example, the classification result being indeed positive (Equation 11). These metrics were chosen since we did not use balanced accuracy to evaluate our method.

$$Specificity = TN/(TN+FP) \quad (10)$$

$$Sensitivity=TP/(TP+FN) \quad (11)$$

To correctly train our network considering the action of falling, since the entire positive video can contain many other different actions, we cropped only the moment in which a fall occurred. Therefore, so that the network was not trained with what we consider a negative situation (walking, sitting, running, etc), we narrowed the training videos to only the period of frames containing a fall. This reduced the training set, however, it was a way to ensure that the training was not corrupted with false cases.

Considering that the training data was diminished and the not falling cases contained a lot more data, to try to balance this situation we applied a random downsampling of the dataset's negative class to match the size of the positive class. Later on, experiments were conducted considering an 80% training set and 20% testing set.

The results for our multi-stream for all of the descriptor combinations for the FDD dataset concerning the different ensemble methods explained in Section 4.4 can be seen in Tables 1, 2 and 3, respectively.

The best result obtained for the FDD dataset shown in Tables 1, 2 and 3 were yielded upon experiments with a five fold cross-validation, learning rate of 10^{-6} , a mini-batch of 1024, 500 epochs and the batches were normalized. It was possible to see that the use of an SVM during the ensemble technique was beneficial for the classification of the FDD dataset. Considering the average and threshold (Table 1) as the baseline of our method, we are able to observe that the use of the SVM

both with the average (Table 2) and continuous values (Table 3) had an improvement concerning the combination of weaker features such as the visual rhythm and the pose estimation. Both the results for Table 2 and 3 were slightly similar, however, when using an SVM with the continuous values, the sensitivity was enhanced by almost 5% in a three stream combination.

Table 1. Results for the FDD dataset with the average and threshold.

Methods	FDD avg		
	Sensitivity	Specificity	Accuracy
Multi-Stream (OF+PE+VR)	42.00%	92.00%	96.80%
Multi-Stream (OF+PE)	85.00%	99.00%	98.66%
Multi-Stream (OF+VR)	23.00%	100.00%	95.73%
Multi-Stream (PE+VR)	29.00%	100.00%	96.09%
Single-Stream (OF)	69.00%	90.00%	98.22%
Single-Stream (PE)	100.00%	25.00%	29.21%
Single-Stream (VR)	13.00%	100.00%	95.20%

Table 2. Results for the FDD dataset with the average and SVM.

Methods	FDD SVM/avg		
	Sensitivity	Specificity	Accuracy
Multi-Stream (OF+PE+VR)	74.00%	100.0%	98.57%
Multi-Stream (OF+PE)	84.00%	100.0%	98.84%
Multi-Stream (OF+VR)	71.00%	100.00%	98.40%
Multi-Stream (PE+VR)	32.00%	100.00%	98.49%
Single-Stream (OF)	77.00%	100.0%	98.22%
Single-Stream (PE)	100.00%	24.00%	28.24%
Single-Stream (VR)	26.00%	100.00%	95.91%

Results for the URFD dataset can be seen in Tables 4, 5 and 6. Similar to the FDD results, the experiments were conducted with a five fold cross-validation, a mini-batch of 1024, learning rate of 10^{-6} , 500 epochs and normalized batches. We have observed that the use of a smaller learning rate helped to improve the results of both of the URFD and FDD datasets that had a smaller amount of positive cases to use while in training. In addition, the accuracy values also had an improvement with the use of an SVM in the classification. The enhancement was made upon sensitivity values as well, as it can be seen that Table 6 had an 11% growth compared to Table 5 in a three-stream case combination.

Table 3. Results for the FDD dataset with an SVM.

Methods	FDD SVM		
	Sensitivity	Specificity	Accuracy
Multi-Stream (OF+PE+VR)	82.00%	99.81%	98.84%
Multi-Stream (OF+PE)	82.00%	99.43%	98.49%
Multi-Stream (OF+VR)	77.00%	99.90%	98.66%
Multi-Stream (PE+VR)	24.00%	100.0%	95.82%
Single-Stream (OF)	77.00%	99.71%	98.49%
Single-Stream (PE)	100.0%	24.00%	28.00%
Single-Stream (VR)	26.00%	100.0%	95.00%

Table 4. Results for the URFD dataset with the average and threshold.

Methods	URFD avg		
	Sensitivity	Specificity	Accuracy
Multi-Stream (OF+PE+VR)	42.00%	100.0%	96.82%
Multi-Stream (OF+PE)	84.00%	99.0%	98.40%
Multi-Stream (OF+VR)	20.00%	100.00%	95.59%
Multi-Stream (PE+VR)	30.00%	100.0%	96.10%
Single-Stream (OF)	61.00%	100.0%	97.86%
Single-Stream (PE)	98.00%	26.00%	29.75%
Single-Stream (VR)	13.00%	100.0%	95.11%

The Multicam dataset was the largest dataset used during the experiments, its results can be seen in Tables 7, 8 and 9. As expected based on the same pattern of the results yielded by the previous tested datasets, Table 9 both sensitivity and specificity values had improvements compared to the other tables when the

ensemble was assisted by the SVM. These results were obtained also using a five fold cross-validation, a mini-batch of 1024, learning rate of 10^{-3} , normalized batches and 1000 epochs. It is possible to observe that Table 9 had the best results, however when dealing with a dataset with a larger amount of data Tables 7 and 8 had similar values.

Table 5. Results for the URFD dataset with the average and SVM.

Methods	URFD SVM/avg		
	Sensitivity	Specificity	Accuracy
Multi-Stream (OF+PE+VR)	66.00%	100.0%	98.13%
Multi-Stream (OF+PE)	71.00%	100.0%	98.40%
Multi-Stream (OF+VR)	61.00%	100.00%	97.86%
Multi-Stream (PE+VR)	23.00%	100.0%	95.73%
Single-Stream (OF)	63.00%	100.0%	97.95%
Single-Stream (PE)	98.00%	26.00%	30.37%
Single-Stream (VR)	19.00%	100.0%	95.55%

Table 6. Results for the URFD dataset with an SVM.

Methods	URFD SVM		
	Sensitivity	Specificity	Accuracy
Multi-Stream (OF+PE+VR)	77.00%	100.0%	98.75%
Multi-Stream (OF+PE)	81.00%	100.0%	98.84%
Multi-Stream (OF+VR)	61.00%	100.00%	97.86%
Multi-Stream (PE+VR)	19.00%	100.0%	95.55%
Single-Stream (OF)	76.00%	100.0%	98.57%
Single-Stream (PE)	98.00%	24.00%	28.33%
Single-Stream (VR)	24.00%	100.0%	95.82%

As it can be seen the accuracy values did not necessarily increase if the number of streams increased. However, we believe that some high-level features can be appropriate to detect specific actions, for example the pose estimation descriptor has a better detection rate when dealing with falling situations while the visual rhythm with negative cases for some of the datasets. Both of these features assisted in the growth of specificity and sensitivity values for the optical flow descriptor.

Table 7. Results for the Multicam dataset with the average and threshold.

Methods	Multicam avg		
	Sensitivity	Specificity	Accuracy
Multi-Stream (OF+PE+VR)	84.00%	92.00%	88.33%
Multi-Stream (OF+PE)	62.00%	76.00%	70.45%
Multi-Stream (OF+VR)	89.00%	95.00%	92.52%
Multi-Stream (PE+VR)	92.00%	82.00%	86.56%
Single-Stream (OF)	58.00%	90.00%	76.35%
Single-Stream (PE)	27.00%	62.00%	47.17%
Single-Stream (VR)	91.00%	92.00%	91.71%

Table 8. Results for the Multicam dataset with the average and SVM.

Methods	Multicam SVM/avg		
	Sensitivity	Specificity	Accuracy
Multi-Stream (OF+PE+VR)	77.00%	93.00%	86.21%
Multi-Stream (OF+PE)	58.00%	86.00%	73.65%
Multi-Stream (OF+VR)	74.00%	98.00%	87.76%
Multi-Stream (PE+VR)	88.00%	90.00%	89.50%
Single-Stream (OF)	55.00%	91.00%	75.79%
Single-Stream (PE)	25.00%	68.00%	49.58%
Single-Stream (VR)	91.00%	92.00%	91.69%

In a general applicability, our method outputs as interesting results as some of the methods found in the literature which used the same datasets (Tables 10, 11 and 12). However, we are not able to directly compare the methods of the literature to our approach, since we do not know how the dataset was manipulated. Nonetheless, our method had better specificity and, in some cases, accuracy results when compared to the others.

In general, the small amount of data provided by these datasets, considering what we use for training, made it difficult to maintain an above 95% accuracy for all of the cases. It is possible to observe that the dataset that had a larger amount of information, Multicam, yielded more stable results compared to FDD and URFD. Compared to Carneiro et al. [1], our method even though not using the RGB information was still able to output compatible results.

Table 9. Results for the Multicam dataset with an SVM.

Methods	Multicam SVM		
	Sensitivity	Specificity	Accuracy
Multi-Stream (OF+PE+VR)	90.00%	94.00%	92.33%
Multi-Stream (OF+PE)	58.00%	90.00%	76.48%
Multi-Stream (OF+VR)	91.00%	93.00%	92.21%
Multi-Stream (PE+VR)	91.00%	92.00%	91.64%
Single-Stream (OF)	55.00%	91.00%	75.79%
Single-Stream (PE)	25.00%	68.00%	49.58%
Single-Stream (VR)	91.00%	92.00%	91.69%

Table 10. FDD comparison with other methods.

Methods	FDD comparison		
	Sensitivity	Specificity	Accuracy
Multi-Stream (OF+PE+VR)	82.00%	99.81%	98.84%
Nunez-Marcos et al.[13]	99.00%	97.00%	97.00%
Zerrouki and Houacine [11]	-	-	97.02%
Carneiro et al. [1]	99.90%	98.32%	98.43%

Table 11. URFD comparison with other methods.

Methods	URFD comparison		
	Sensitivity	Specificity	Accuracy
Multi-Stream (OF+PE+VR)	77.00%	100.0%	98.75%
Nunez-Marcos et al.[13]	100.0%	92.00%	95.00%
Zerrouki and Houacine [11]	-	-	96.88%
Kwolek and Kepski [12]	100.0%	92.50%	95.71%
Carneiro et al. [1]	100.0%	98.61%	98.77%

Table 12. Multicam comparison with other methods.

Methods	Multicam comparison		
	Sensitivity	Specificity	Accuracy
Multi-Stream (OF+PE+VR)	90.00%	94.00%	92.33%
Nunez-Marcos et al.[13]	99.00%	96.00%	-

5.2. Qualitative Results

As expected, the method had some troubles when dealing with situations that had similar movement as falling actions. The videos that were not classified correctly were related to abruptly sitting actions, squatting or even lying down (Figure 8).



Fig. 8. Misleading fall detection squatting frame example.

6. CONCLUSIONS AND FUTURE WORK

In this work, we proposed and evaluated a multi-stream learning model based on convolutional neural networks to cope with a falling classification problem. Therefore, our approach consisted in extracting hand-crafted high-level features (optical flow, visual rhythm, and pose estimation) from public data set videos and using each one as an input to a distinct VGG-16 classifier. In addition to the feature combination, we also studied the best of three ensemble techniques to cope with our binary classification problem.

We believe that a multi-stream model can assist in classification since an outlier result from one of the streams can be corrected based on the other. Using high-level features also assisted in covering unnecessary information from the video frames such as the background and other unimportant details. It was also possible to observe that the SVM assisted in the balance and increase of the sensitivity and specificity metrics when used in the ensemble. In addition, compared to previous work we were able to maintain accuracy even though not using the RGB frame information itself.

For future work, we intend to continue investigating relevant high-level features used for fall detection. In addition, the studies will be conducted to test better parameters for the individual learners and use other architectures for each stream. Finally, we will also investigate the use of contexts other than falling to observe if this learning method is able to maintain its accuracy and general applicability.

7. ACKNOWLEDGMENTS

The current archival periodical article is based on the conference presentation [1]. The authors thank CAPES, FAPESP (grants #2014/12236-1 and #2017/12646-3), CNPq (grants #305169/2015-7, #421521/2016-3 and #307062/2016-3) for the financial support, as well as Semantix Brasil for the infrastructure and support provided during the development of the present work.

8. REFERENCES

- [1] S. Carneiro, G. Silva, G. Leite, R. Moreno, S. Guimaraes, H. Pedrini, "Multi-Stream Deep Convolutional Network Using High-Level Features Applied to Fall Detection in Video Sequences", Proceedings of the 26th International Conference on Systems, Signals and Image Processing, Osijek, Croatia, 5-7 June 2019, pp. 293–298.
- [2] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", arXiv:1409.1556, 2014.
- [3] C.-W. Lin, Z.-H. Ling, Y.-C. Chang, C. J. Kuo, "Compressed-Domain Fall Incident Detection for Intelligent Home Surveillance", Proceedings of the International Symposium on Circuits and Systems, Kobe, Japan, 23-26 May 2005, pp. 3781–3784.
- [4] D. T. Anderson, J. M. Keller, M. Skubic, X. Chen, Z. He, "Recognizing Falls from Silhouettes", Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, USA, 30 August - 3 September 2006, pp. 6388–6391.
- [5] H. Qian, J. Zhou, Y. Mao, Y. Yuan, "Recognizing Human Actions from Silhouettes Described with Weighted Distance Metric and Kinematics", Multimedia Tools and Applications, Vol. 76, No. 21, 2017, pp. 21889–21910.
- [6] M. Yu, Syed . Naqvi, J. Chambers, "Fall Detection in the Elderly by Head Tracking", Proceedings of the 5th Workshop on Statistical Signal Processing, Cardiff, UK, 31 August - 3 September 2009, pp. 357–360.
- [7] M. Yu, S. M. Naqvi, J. Chambers, "A Robust Fall Detection System for the Elderly in a Smart Room", Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing, Dallas, TX, USA, 14-19 March 2010, pp. 1666–1669.

- [8] C. Rougier, J. Meunier, A. St-Arnaud, J. Rousseau, "Robust Video Surveillance for Fall Detection based on Human Shape Deformation", *Proceedings of the IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 21, No. 5, 2011, pp. 611–622.
- [9] B.-S. Lin, J.-S. Su, H. Chen, C. Y. Jan, "A Fall Detection System based on Human Body Silhouette", *Proceedings of the 9th International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, Beijing, China, 16-18 October 2013, pp. 49–52.
- [10] D. P. Kumar, Y. Yun, I. Yu-Hua Gu, "Fall Detection in RGB-D Videos by Combining Shape and Motion Features", *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Shanghai, China, 20-25 March 2016, pp. 1337–1341.
- [11] N. Zerrouki, F. Harrou, A. Houacine, Y. Sun, "Fall Detection using Supervised Machine Learning Algorithms: A Comparative Study", *Proceedings of the 8th International Conference on Modelling, Identification and Control*, Algiers, Algeria, 15-17 November 2016, pp. 665–670.
- [12] B. Kwolek, M. Kepski, "Improving Fall Detection by the Use of Depth Sensor and Accelerometer", *Neurocomputing*, Vol. 168, 2015, pp. 637–645.
- [13] A. Nunez-Marcos, G. Azkune, I. Arganda-Carreras, "Vision-based Fall Detection with Convolutional Neural Networks", *Wireless Communications and Mobile Computing*, Vol. 2017, 2017.
- [14] E. Akağunduz, M. Aslan, A. Senğur, H. Wang, M. C. Ince, "Silhouette Orientation Volumes for Efficient Fall Detection in Depth Videos", *IEEE Journal of Biomedical and Health Informatics*, Vol. 21, No. 3, 2017, pp. 756–763.
- [15] L. Anishchenko, "Machine Learning in Video Surveillance for Fall Detection", *Proceedings of the Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology*, Yekaterinburg, Russia, 7-8 May 2018, pp. 99–102.
- [16] N. Zerrouki, A. Houacine, "Combined Curvelets and Hidden Markov Models for Human Fall Detection", *Multimedia Tools and Applications*, Vol. 77, No. 5, 2018, pp. 6405–6424.
- [17] A. Lowhur, M. C. Chuah, "Dense Optical Flow based Emotion Recognition Classifier", *Proceedings of the 12th International Conference on Mobile Ad Hoc and Sensor Systems*, Dallas, TX, USA, 19-22 October 2015, pp. 573–578.
- [18] B. S. Torres, H. Pedrini, "Detection of Complex Video Events through Visual Rhythm", *The Visual Computer*, Vol. 34, No. 2, 2018, pp. 145–165.
- [19] T. Moreira, D. Menotti, H. Pedrini, "First-Person Action Recognition Through Visual Rhythm Texture Description", *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, New Orleans, LA, USA, 5-9 March 2017, pp. 1–4.
- [20] F. B. Valio, H. Pedrini, N. J. Leite, "Fast Rotation-Invariant Video Caption Detection based on Visual Rhythm", *Proceedings of the Iberoamerican Congress on Pattern Recognition*, Pucón, Chile, 15-18 November 2011, pp. 157–164.
- [21] S. J. F. Guimaraes, M. Couprie, A. A. Araújo, and N. J. Leite, "Video Segmentation based on 2D Image Analysis", *Pattern Recognition Letters*, Vol. 24, No. 7, 2003, pp. 947–957.
- [22] S. J. F. Guimaraes, A. A. Araújo, M. Couprie, N. J. Leite, "Video Fade Detection by Discrete Line Identification," *Proceedings of the Object Recognition Supported by User Interaction for Service Robots*, Quebec City, Quebec, Canada, Canada, 11-15 August 2002, pp. 1013–1016.
- [23] Z. Cao, T. Simon, S.-E. Wei, Y. Sheikh, "Real time Multi-Person 2D Pose Estimation using Part Affinity Fields", *arXiv preprint arXiv:1611.08050*, 2016.
- [24] T. G. Dietterich, "Ensemble Methods in Machine Learning", *Proceedings of the International Workshop on Multiple Classifier Systems*, Springer, Cagliari, Italy, 21-23 June 2000, pp. 1–15.
- [25] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, "Imagenet: A Large-Scale Hierarchical Image Database", *Proceedings of the Conference on Computer Vision and Pattern Recognition*, Miami, FL, USA, 20-25 June 2009, pp. 248–255.
- [26] K. Soomro, A. R. Zamir, M. Shah, "UCF101: A Dataset of 101 Human Actions Classes from Videos in the Wild", *arXiv preprint arXiv:1212.0402*, 2012.

- [27] B. Kwolek, M. Kepski, "Human Fall Detection on Embedded Platform using Depth Maps and Wireless Accelerometer", *Computer Methods and Programs in Biomedicine*, Vol. 117, No. 3, 2014, pp. 489–501.
- [28] H. A. Maia, D. T. Concha, H. Pedrini, H. Tacon, A. S. Brito, H. L. Chaves, M. B. Vieira, S. M. Villela, "Action Recognition in Videos Using Multi-Stream Convolutional Neural Networks", *Deep Learning Applications*, Springer International Publishing, 2020, pp. 95–111.
- [29] R. Quispe, D. Ttito, A. Rivera, H. Pedrini, "Multi-Stream Networks and Ground-Truth Generation for Crowd Counting", *International Journal of Electrical and Computer Engineering Systems*, Vol. 11, No. 1, 2020, pp. 25–33.
- [30] H. Tacon, A. S. Brito, H. L. Chaves, M. B. Vieira, S. M. Villela, H. A. Maia, D. T. Concha, H. Pedrini, "Multi-Stream Architecture with Symmetric Extended Visual Rhythms for Deep Learning Human Action Recognition", *Proceedings of the 15th International Conference on Computer Vision Theory and Applications*, Valletta, Malta, February 27-29, 2020, pp. 351–358.

The Effect of Space-filling Curves on the Efficiency of Hand Gesture Recognition Based on sEMG Signals

Original Scientific Paper

Panagiotis Tsinganos

University of Patras, Department of Electrical and Computer Engineering, 26504 Patras, Greece
Vrije Universiteit Brussel, Department of Electronics and Informatics, 1050 Brussels, Belgium
panagiotis.tsinganos@ece.upatras.gr

Bruno Cornelis

Vrije Universiteit Brussel, Department of Electronics and Informatics, 1050 Brussels, Belgium
IMEC, 3001 Leuven, Belgium
bcorneli@etrovub.be

Jan Cornelis

Vrije Universiteit Brussel, Department of Electronics and Informatics, 1050 Brussels, Belgium
jpcornel@etrovub.be

Bart Jansen

Vrije Universiteit Brussel, Department of Electronics and Informatics, 1050 Brussels, Belgium
IMEC, 3001 Leuven, Belgium
bjansen@etrovub.be

Athanassios Skodras

University of Patras, Department of Electrical and Computer Engineering,
26504 Patras, Greece
skodras@ece.upatras.gr

Abstract – Over the past few years, Deep learning (DL) has revolutionized the field of data analysis. Not only are the algorithmic paradigms changed, but also the performance in various classification and prediction tasks has been significantly improved with respect to the state-of-the-art, especially in the area of computer vision. The progress made in computer vision has produced a spillover in many other domains, such as biomedical engineering. Some recent works are directed towards surface electromyography (sEMG) based hand gesture recognition, often addressed as an image classification problem and solved using tools such as Convolutional Neural Networks (CNN). This paper extends our previous work on the application of the Hilbert space-filling curve for the generation of image representations from multi-electrode sEMG signals, by investigating how the Hilbert curve compares to the Peano- and Z-order space-filling curves. The proposed space-filling mapping methods are evaluated on a variety of network architectures and in some cases yield a classification improvement of at least 3%, when used to structure the inputs before feeding them into the original network architectures.

Keywords – classification, CNN, Deep Learning, electromyography, hand gesture recognition, Hilbert curve, Peano curve, sEMG, space-filling curve, Z-order curve

1. INTRODUCTION

Hand gesture recognition finds many applications in human computer interaction [1], sign language recognition [2], prosthesis control [3] and gaming for rehabilitation [4, 5]. The electrical signals generated by the forearm muscles during the execution of hand gestures provides useful information about muscle activity and hand motion [6]. These signals can be recorded by means of surface electromyography (sEMG) sensors.

The type of hand motions and gestures has tradition-

ally been determined from sEMG data using a combination of (hand-crafted) feature extractors and Machine Learning (ML) classifiers. More precisely, the components of an ML-based pattern recognition system include data acquisition, feature extraction, classification, and inference from new data. In the hand gesture recognition case, the electrodes (sensors) that acquire the sEMG signals, are attached to the arm and/or forearm. Then, features such as Root Mean Square (RMS), variance, zero crossings and frequency coefficients are extracted from the sEMG signals. Finally, these features

are fed into classifiers like k-Nearest Neighbours (k-NN), Support Vector Machine (SVM), Multi-Layer Perceptron (MLP) or Random Forests [7].

The success of Deep Learning (DL) models in solving the problem of sEMG-based hand gesture recognition has been well-documented in numerous studies. In these works, a usual approach is to represent sEMG data as images and feed them into a Convolutional Neural Network (CNN) that outputs a probability that an sEMG signal corresponds a certain gesture class. In a typical CNN architecture, the output label is obtained from a sequence of convolutional and pooling layers followed by fully connected (i.e. dense) layers and a softmax activation layer. Consequently, CNN models transform the input image layer by layer, from its pixel values to the final classification label.

Although an architecture based on Recurrent Neural Networks (RNN) would seem more appropriate for processing time-series signals, little research has been done on this. The problem might be that Long Short-Term Memory (LSTM) networks, the main type of RNN models, require many parameters which cannot be trained efficiently with the limited sEMG datasets available. In addition, recent work [8, 9] has shown that CNN can outperform RNN in a wide range of sequence problems.

There has been considerable progress in the use of CNNs for feature extraction, image classification and particularly for combined feature extraction and classification within the same network. On the other hand, converting time-series into image-like structures that can be used as inputs to CNN models is not a simple task. Methods proposed in the open literature include the partitioning of multi-channel signals using windows and the application of 2D transformations such as the Fourier and Wavelet Transforms.

The current archival periodical article is based on the conference presentation [10], in which we describe a representation method for sEMG signals using the Hilbert curve to obtain images. Now, we investigate the application of several space-filling curves, namely the Hilbert, Peano, and Z-order curves, to represent sEMG signals as images that can be classified by CNNs. Compared to [10], the added contribution presented in this paper is the evaluation of all these types of image representation methods and their performance comparison when applied on the problem of hand gesture recognition using CNNs.

The rest of the paper is structured as follows. Section 2 contains a literature review of the methodologies used for hand gesture recognition. The details of the proposed methods and the CNN architectures applied in our experiments are given in Section 3. The experiments performed for the evaluation of the models are given in Section 4, while the results and a discussion of these results are presented in Section 5. Finally, Section 6 concludes by summarizing the outcomes of this work.

2. RELATED WORK

Hand gesture recognition based on sEMG has been investigated by both typical ML approaches and DL practices. In the case of ML methods, the first approach [11] for the classification of four gestures utilizes time-domain features from two-electrode sEMG signals. An accuracy of 97% in categorizing three types of grasps is achieved in the work of [12] by using the RMS feature of seven electrodes as input to an SVM classifier. The authors of [13, 14, 15] evaluate a large amount of EMG features on a selection of different classifiers for the classification of 52 gestures from the Ninapro reference dataset [13, 16]. The highest performance, an accuracy of 75%, was achieved by a Random Forest classifier using a combination of statistical and frequency domain features.

In the case of DL methods, a large body of research results has been recently developed. The CNN defined in [17] for the recognition of six common gestures resulted in improved accuracy compared to the one obtained by applying an SVM classifier. In [18], a performance comparable to ML approaches is achieved by a model consisting of convolutional and average pooling layers. In [19], we developed techniques to increase the accuracy of the simple model, presented in [18]. Choosing max pooling and inserting dropout layers to reduce overfitting as suggested in [20], allowed us to obtain a 3% increase in accuracy (i.e. from 67% to 70%). The authors of [21, 22] follow a different approach not only in the choice of the model architecture but also in the recording of EMG signals. Their work is based on high-density electrode arrays, which is considered an effective approach in myoelectric control as described in [23, 24, 25]. Using instantaneous EMG images, the CNN model of [21] correctly classifies eight hand movements achieving an accuracy of 89%. The multi-stream CNN described in [22] achieves an accuracy of 85% on the Ninapro dataset, which consists of 52 movements.

Research that utilizes different DL approaches has been carried out as well. The work presented in [26] follows a model adaptation strategy that applies adaptive batch normalization (AdaBN) [27] for updating the parameters of the normalization layers. The use of weighted connections between a source and a target network is presented in [28]. Data augmentation practices for sEMG signals are developed and evaluated in [28]. More recent work in this topic is presented in [29] where magnitude warping and wavelet decomposition showed considerable improvement in classification accuracy.

The use of fractal curves, and in particular the Hilbert curve, for designing alternative image and signal representations is long known, but never studied for sEMG signals. In [30] and [31], the properties of the Hilbert curve have been exploited to convert mammographic images to 1D vectors. In combination with a set of appropriate features, this helped in detecting breast can-

cer. In [32] a similar dimensionality reduction approach is followed. The mapping of 3D data into 2D and 1D representations facilitates the classification of 3D structural data by CNNs. Compared to the direct processing of raw data, such an approach reduces training time and can be used in cases of data of an arbitrary number of channels. The sequence of the extracted image patches is of importance for the detection of image-forgeries by LSTM-based models [33]. The Hilbert curve is employed to determine the order of the image patches fed into the LSTM, thus preserving spatial locality.

The Hilbert curve can be also used in the inverse problem, i.e. transforming 1D data in 2D images. The authors of [34] found that long-term interactions between regions of the DNA sequence are important for its classification. Thus, instead of very deep networks or larger filters, the Hilbert curve was employed to map the DNA sequence into an image such that proximal elements stay close, and the distance between distant elements is reduced.

Besides the more popular Hilbert curve, also Z-order curves have been used in representation problems, though less frequently. For example, in [35] the Z-order curve is chosen because of its good balance between locality preservation and computational complexity to map the neighbourhood around a point in a 3D point cloud into a 1D sequence. These sequences are fed into a CNN to predict the displacement between the current and the next point. The authors of [36] represent the 4D coordinates of a crystal structure as a 1D feature vector using the Z-order curve. An MLP predictor takes as input these representations and estimates the energy of the organic molecular structure.

3. PROPOSED DATA REPRESENTATION METHODS

3.1. Space-filling curves

Space-filling curves, such as the Peano, Hilbert, and Z-order (Morton) curves, have found applications in various domains, such as database access, data compression, and image processing. The key property of

space-filling curves is that they constitute a mapping between a multidimensional space and a lower dimensional space while, in general, they preserve locality between the data points.

The Hilbert curve or Hilbert space-filling curve was first described by the German mathematician David Hilbert in 1891. The main property of this continuous fractal space-filling curve is the superiority in preserving locality compared to alternative curves [37, 38]. The Hilbert curve can be constructed recursively. Firstly, the 2D plane is divided into four quadrants that are traversed according to a fundamental pattern as shown in Fig. 1-(a). In each subsequent iteration, all existing subsquares are subdivided into four smaller subsquares. These four subsquares are connected by a pattern obtained by rotation and/or reflection of the fundamental pattern. Fig. 1-(a) visualizes Hilbert curve traversals of the 2D space after the first iterations, where the numbers are the index within the 1D sequence that is mapped to the specific pixel of the 2D image.

The Peano curve is the first example of a space-filling curve to be discovered, by Giuseppe Peano in 1890. To construct the Peano curve the 2D plane is initially divided into nine squares (i.e. a 3×3 grid) that are traversed according to a fundamental pattern. From each iteration to the next, all existing subsquares are subdivided into nine smaller subsquares connected as shown in Fig. 1-(b).

The Z-order (also known as Morton curve) is another locality preserving space-filling curve named after Guy Macdonald Morton, who first applied the order to file sequencing in 1966. Its basic pattern looks like the letter 'Z'. The curve is constructed as shown in Fig. 1-(c).

3.2. sEMG representation

Fractal curves are used to transform multi-channel sEMG signals into 2D image representations. The sEMG signals are recorded by K electrodes (channels), while performing a hand gesture. These are organized into small segments of N samples. Such an arrangement results in sEMG data of size $N \times K$. The fractal curve map-

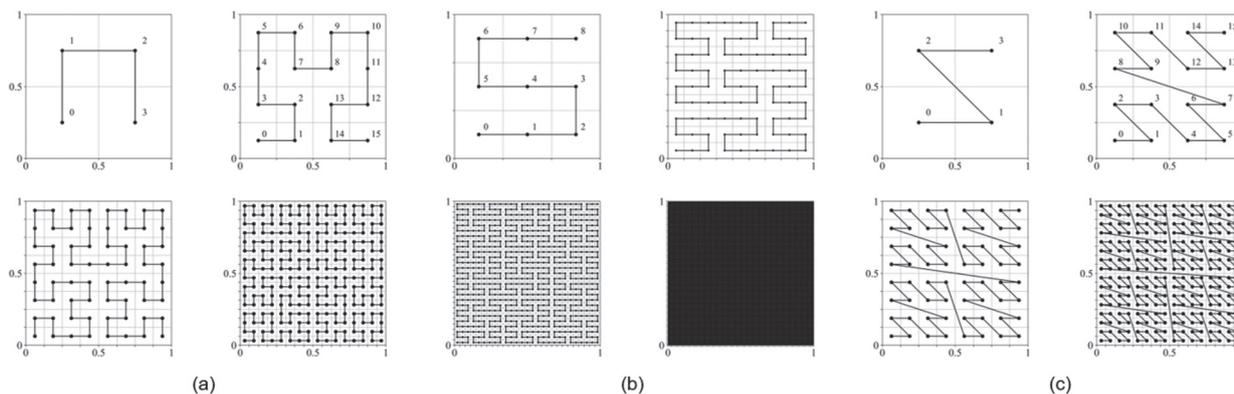


Fig. 1. First iterations of the Hilbert (a), Peano (b) and Z-order (c) curves.

ping can be applied in two ways, either (i) across the time dimension, i.e. map each time sequence of sEMG channel into a 2D image, or (ii) across the sEMG channels, i.e. map the values of the sEMG channels of each time instant into a 2D image (Fig. 2).

The application of the Hilbert and Z-order mappings across the time dimension is as follows. Given a single-electrode sEMG sequence of length N , a 2D representation of $M \times M$ size is produced, where $N=M^2$ and M is a power of two, i.e. $M=2^n$. In the case of K sEMG electrodes, this process is repeated for every electrode, and the outputs are stacked into a K -channel image, resulting in an $M \times M \times K$ image. For example, an sEMG signal of 10 electrodes and 64 samples is mapped into an $8 \times 8 \times 10$ image. The application of the Peano curve is similar, but M should be a power of three, i.e. $M=3^n$. It is important to note that sequence segments of length smaller than M^2 , i.e. $(M-1)^2 < N < M^2$, can be used as well, however, in that case the final image has to be filled out with zeros up to a length of M^2 .

In the second option, the curve mapping is applied across the sEMG electrodes. The number of sEMG electrodes should be a square number, i.e. $K=M^2$, where $M=2^n$ (for the Hilbert and Z-order curves) or $M=3^n$ (for the Peano curve). At every time instant of the sequence the set of K electrode values is mapped into an $M \times M$ image. Thus the dimensions of the Hilbert curve representation of the $N \times K$ segment will be $M \times M \times N$. For example, an sEMG segment of 16 electrodes with 20 samples is mapped into a $4 \times 4 \times 20$ image. As in the previous case, the image will be zero-padded, if there are less electrodes than M^2 (Fig. 2-c).

The computation of the mapping from 1D to 2D using these space-filling curves requires only bitwise operations based on Gray codes performed in $O(1)$ time. For a given $M \times M$ grid, calculating the coordinates for a single data point requires $O(\log_2 M)$ repetitions [39]. Since the mapping is the same for all images, it is computed only once and then used as a look-up table. Therefore, the computational overhead is very small compared to training the CNN.

A common approach in image applications is to use 1D (grayscale) or 3D (RGB) images as inputs to CNNs. In our approach, the dimensionality of the image corresponds either to the number of electrodes or to the sEMG segment duration.

4. EXPERIMENTS

4.1. The dataset

For the evaluation of the proposed models the first dataset of the Ninapro database [13] was selected. This consists of sEMG recordings of 27 healthy subjects that repeat 52 gestures 10 times. The hand movements can be grouped into three categories: (i) basic finger movements, (ii) isometric, isotonic hand configurations and wrist movements, and (iii) grasps and functional movements. EMG signals are measured using 10 electrodes. Eight of them are equally spaced around the forearm. The remaining electrodes are placed on the main activity spots of the large flexor and extensor muscles of the forearm [13].

As in previous studies involving the Ninapro dataset, the sEMG signals are pre-processed by a low-pass

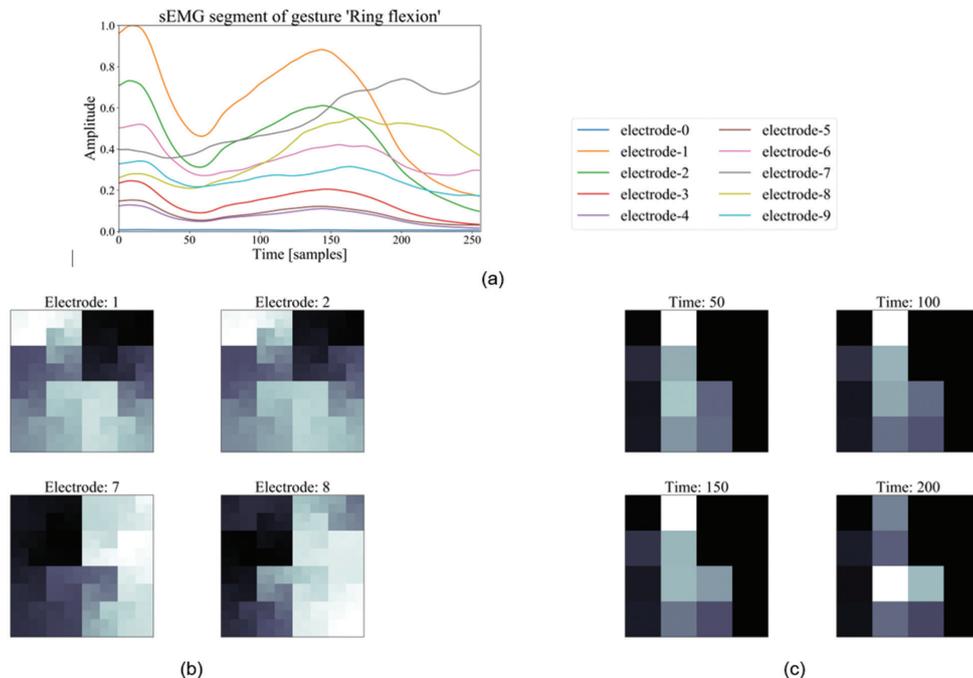


Fig. 2. The application of the Hilbert curve mapping to sEMG data [10]. (a) A 256-samples segment of sEMG signal, (b) the Hilbert representation (16×16) across time of electrodes 1, 2, 6, 8 and (c) the Hilbert representation (4×4) across electrodes at time instants 50, 100, 150, 200. In (c) there are less electrodes than pixel dimensions, thus the last six pixels on the right of the images are zero-filled.

filter [18, 19, 21]. Augmentation of the training data is achieved by duplicating the signals of each repetition and adding Gaussian noise of 25dB signal to noise ratio (SNR). Magnitude-warping is also used for sEMG signal

augmentation [40, 41]. As a last step, sEMG signals from the $K = 10$ channels are segmented into overlapping windows of length N with a step of 50ms and are organized into $N \times 10$ arrays.

Table 1. CNN model architectures used in this work [10]

AtzoriNet [18, 19]	VGGNet [42]	DenseNet [43]	SqueezeNet [44]
CONV(32, 1×10), ReLU,	CONV(16, 3×3), BN, ReLU,	CONV(16, 3×3)	CONV(16, 3×3), BN, ReLU,
CONV(32, 3×3), ReLU,	CONV(32, 3×3), ReLU,	DSBLOCK(4), TRBLOCK(50),	SQBLOCK(4, 16, 32),
POOL(max, 3×3),	CONV(32, 3×3), ReLU,	DSBLOCK(3), TRBLOCK(75),	SQBLOCK(3, 16, 64),
CONV(64, 5×5), ReLU,	POOL(max, 2×2),	ReLU,	CONV(G, 1×1), ReLU,
POOL(max, 3×3),	CONV(64, 3×3), ReLU,	CONV(G, 1×1), ReLU,	GLPOOL(avg),
CONV(64, 5×1), ReLU,	CONV(64, 3×3), ReLU,	GLPOOL(avg),	Softmax
CONV(G, 1×1), Softmax	POOL(max, 2×2),	Softmax	
	GLPOOL(avg),		
	FC(G), Softmax		
		DSBLOCK(n)	SQBLOCK(a, b, c)
		{BN, ReLU, CONV(16, 3×3)} _n	{FIRE(b, c)} _a , POOL(max)
		TRBLOCK(k)	FIRE(b, c)
		BN, CONV(k, 1×1), POOL(max)	CONV(b, 1×1), ReLU,
			CONV(c, 1×1), ReLU + CONV(c, 3×3), ReLU
Parameters	85K	74K	71K
			69K

CONV: convolutional layer, POOL: pooling layer, GLPOOL: global pooling layer, FC: fully-connected layer, BN: batch normalization, { }_k: repetition k times, '+': concatenation
G: number of gesture labels

4.2. Network architectures

A number of neural networks have been used for hand gesture recognition [19]. Apart from those, we also investigate CNN architectures that are typically found in image related tasks, but have not yet been applied to sEMG-based hand gesture recognition, such as VGGNet [42], DenseNet [43], and SqueezeNet [44]. For the comparisons to be fair, an effort is made to keep the number of trainable parameters of the networks approximately equal. The model architectures are shown in Table 1.

4.3. Baseline

As our baseline for comparison, we follow the approach where none of the fractal curve mappings is used. This means that the $N \times 10$ arrays are fed into the CNN models as single-channel images. For the window length N , we experiment with two different values: 16 and 64 samples. The reason for choosing $N = 16$ is that for a real-time application the window size should be as small as possible [11, 45]. We also report results for a bigger window, i.e. $N = 64$, because of the highest accuracy that was achieved during the validation experimentation (Fig. 3). The explanation for the trend in Fig. 3, is that with longer segments ($N > 64$) one can generate less training examples compared to smaller N values. Therefore, the CNN models tend to overfit. On the other hand, too small segments do not contain enough information for the classification task.

4.4. Mapping across time (xxxxTime)

Regarding the space-filling curve mapping across the time dimension (xxxxTime), the $N \times 10$ segments are organized into $M \times M \times 10$ images. In the case of Hilbert and Z-order mappings, for N values equal to 16 and 64 the resulting image sizes are $4 \times 4 \times 10$ and $8 \times 8 \times 10$, respectively, while for the Peano curve the image sizes are $4 \times 6 \times 10$ (zero-padded and cropped) and $9 \times 9 \times 10$ (zero-padded).

4.5. Mapping across electrodes (xxxxElect)

The mapping across the sEMG channel dimension (xxxxElect) is performed in a similar fashion. Given the number of channels $K = 10$, the $N \times 10$ segments are organized into images with dimensions $4 \times 4 \times N$. The pixels corresponding to the last six positions that the three fractal curves traverse are set to zero.

In this approach we retain the spatial resolution constant due to the small number of available electrodes. Regarding the window length, we only experimented with $N = 16$, since a longer segment would generate very deep image representations that in turn would increase the number of parameters in the first convolutional layer increasing the probability of overfitting.

4.6. Model hyper-parameters

All networks were trained using stochastic gradient descent for 60 epochs with an initial learning rate of 0.1

halved every 15 epochs, and a batch size of 1024. To avoid overfitting the networks due to the small training set, dropout layers were appended after each convolutional layer with a forget rate of 0.3. In addition, weight decay regularization with a value of 0.0005 was applied to all convolutional layers. These values were selected after performing a grid search on a validation set of ten randomly selected subjects.

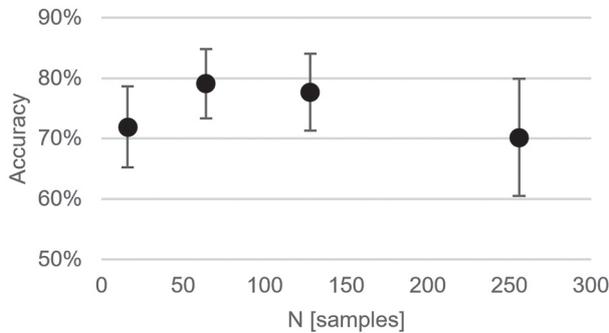


Fig. 3. Validation accuracy vs window size for $N=\{16, 64, 128, 256\}$. Error bars correspond to standard deviation. Highest performance is achieved for $N = 64$.

4.7. Evaluation

The evaluation follows the approach that has been used by other researchers that use the Ninapro dataset [18, 19, 21]. Specifically, a new model is trained for each subject on the data of seven repetitions and is tested on the remaining three. The performance metrics used are the top-1 and top-3 accuracies, i.e. the accuracy when the highest and any of the 3 highest output probabilities match the expected gesture, as well as the precision and recall values. The average across the 27 subjects in the dataset is reported for each metric. For the statistical comparison of the methods, repeated measures Analysis of Variance (ANOVA) and paired Wilcoxon tests are employed, with the F-value being the test statistic and the p-value the corresponding probability.

5. RESULTS

In the experiments, we evaluate two methods for generating sEMG image representations from multi-channel sEMG signals using three different space-filling curves. The evaluation across four CNN models is shown in Table 2, where an ^{‘*’} denotes a significant difference (Wilcoxon

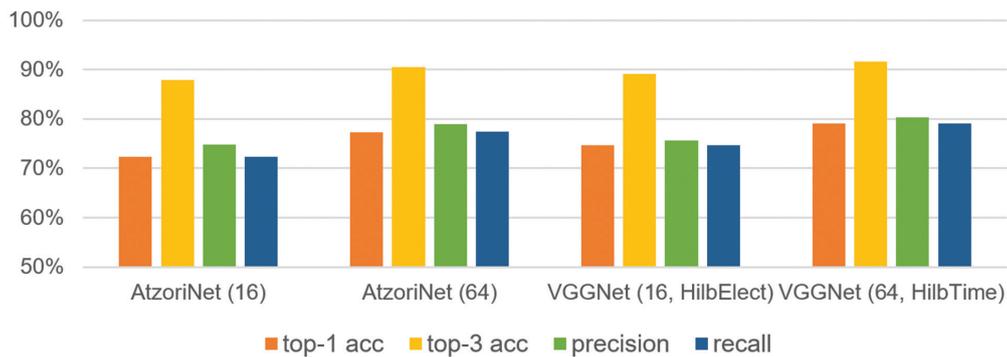


Fig. 4. Comparison of the evaluation metrics for the AtzoriNet[18, 19] and the best performing combinations of representation method and CNN model from Table 2 for $N = 16$ and $N = 64$.

Table 2. Top-1 accuracy metric across sEMG representation method and CNN model. An ^{‘*’} denotes a significant difference ($\alpha = 5\%$) between the baseline and the corresponding space-filling representation. Values in parentheses correspond to the standard deviation.

Segment	Model	Baseline	HilbTime	HilbElect	PeanTime	PeanElect	ZordTime	ZordElect
N = 16	VGGNet	0.7115	0.7192*	0.7469*	0.6957*	0.7346*	0.7045	0.7404*
		(0.0682)	(0.0670)	(0.0653)	(0.0688)	(0.0648)	(0.0685)	(0.0639)
	DenseNet	0.7225	0.7064*	0.7319*	0.6838*	0.7069*	0.6866*	0.7025*
		(0.0639)	(0.0634)	(0.0598)	(0.0689)	(0.0562)	(0.0613)	(0.0597)
	SqueezeNet	0.5611	0.5585	0.5742	0.4062*	0.5934	0.3986*	0.5908
		(0.2342)	(0.1995)	(0.2144)	(0.3017)	(0.1798)	(0.2514)	(0.1816)
N = 64	VGGNet	0.7592	0.7908*	-	0.7527	-	0.7535	-
		(0.0629)	(0.0570)	-	(0.0560)	-	(0.0590)	-
	DenseNet	0.7493	0.7757*	-	0.7544	-	0.7449	-
		(0.0648)	(0.0588)	-	(0.0535)	-	(0.0547)	-
	SqueezeNet	0.6297	0.5188	-	0.3307*	-	0.3281*	-
		(0.1445)	(0.3298)	-	(0.3499)	-	(0.3297)	-

Table 3. Repeated measures ANOVA. On the left, the F and p values of the ANOVA are reported for N = 16 and N = 64. On the right, the p-values of the pairwise comparisons based on the Wilcoxon test are shown. Values above the diagonal correspond to N = 16, and the values below the diagonal are for N = 64. An 'x' denotes an invalid comparison, and an '*' a significant difference ($\alpha = 5\%$).

		HilbTime	PeaNTime	ZordTime	HilbElect	PeaNElect	ZordElect
ANOVA (N = 16)		HilbTime	-	9e-5 *	0.0006 *	x	x
F = 9.6077	p = 8e-8 *	PeaNTime	0.0003 *	-	0.4004	x	x
		ZordTime	0.0010 *	0.6309	-	x	x
ANOVA (N = 64)		HilbElect	x	x	x	-	0.1075
F = 7.7856	p = 0.0011 *	PeaNElect	x	x	x	x	-
		ZordElect	x	x	x	x	-

signed rank test at $\alpha = 5\%$ significance level) between the corresponding representation and the baseline method. Fig. 4 shows the evaluation metrics (top-1 and top-3 accuracies, precision, and recall) for the best performing combination of representation method and CNN model for N = 16 and N = 64. In Table 3, repeated measures ANOVA followed by Wilcoxon signed rank tests assess the significance of the differences between the space-filling curves. On the left, the p-values of the ANOVA are reported, while on the right side of the Table, the pairwise comparisons based on the Wilcoxon test are shown, where the values above the diagonal correspond to p-values for N = 16, and the values below the diagonal are for N = 64. An 'x' denotes an invalid comparison, and an '*' a significant difference ($\alpha = 5\%$).

In general, from Table 2 we see that the Hilbert curve mappings perform always equally well or better compared to the baseline. On the other hand, the image representations of the other two space-filling curves are mostly inferior to the baseline, except for the representations across the electrode dimension (PeaNElect and ZordElect). A comparison between the performances of the CNN architectures, reveals that the VGGNet and the DenseNet yield similar results, whereas the performance of the SqueezeNet is always lower. Considering how much these architectures differ, it is difficult to identify which model components are responsible for this degradation in accuracy performance. The highest top-1 accuracy on the test data is achieved with HilbElect for N = 16 and HilbTime for N = 64, which improves the baseline performance by more than 3%. These representations perform better than the performance of the AtzoriNet architecture, developed specifically for the problem of hand gesture recognition, as shown in Fig. 4.

In Table 3, comparing the top-1 accuracy of the space-filling curves representations shows significant differences for both N = 16 and N = 64. Then, the pairwise comparisons confirm that the sEMG images generated by the Hilbert curve lead to higher performance. In addition, the differences in accuracy between the Peano and the Z-order curves are not significant in general.

For the difference between the representations across time (xxxxTime) and across the electrodes (xxxxElect), we can assume that the former yield better results, since

the dataset used in this work contains data from only ten electrodes. Thus, approximately 1/3 of the pixels of every image correspond to zero padded values, resulting in a limited set of patterns that can be detected by the CNN models compared to the xxxxTime images.

Regarding the advantage of the Hilbert curve in detection accuracy, we can attribute it to the better locality preserving properties between the examined curves. The reason might be that less convolutions are required to extract useful patterns in data, since data are tightly clustered. Therefore, given a network architecture the Hilbert curve image allows for a better utilization of the model parameters.

6. CONCLUSIONS

This work explored the application of the space-filling curves as a means of representing sEMG signals as images for solving the problem of hand gesture recognition. Three curves, i.e. Hilbert, Peano, and Z-order, and two mapping approaches, i.e. traversal across the time dimension and traversal across the electrodes dimension, were evaluated. We made an experimental comparison of three CNN models based on architectures that have been widely used in image processing tasks along with a model optimized for the problem of hand gesture recognition. The results showed that in general, gesture detection accuracy can be improved if sEMG signals are converted to images using fractal curves. The improvement is more evident in the Hilbert curve representations, where the classification accuracy is significantly increased by more than 3% using a VGG-based network. We speculate that this is probably due to the better localization preserving properties of the Hilbert curve. Finally, the differences in accuracy between the Peano and the Z-order curves were not significant in general.

Acknowledgements

The current archival periodical article is based on the conference presentation [10].

The work is supported by the «Andreas Mentzopoulos Scholarships for the University of Patras» and the VUB-UPatras International Joint Research Group on ICT (IJCT).

7. REFERENCES

- [1] S. S. Rautaray, A. Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey", *Artificial Intelligence Review*, Vol. 43, No. 1, 2015, pp. 1–54.
- [2] M. J. Cheok, Z. Omar, M. H. Jaward, "A review of hand gesture and sign language recognition techniques", *International Journal of Machine Learning and Cybernetics*, Vol. 10, No. 1, 2019, pp. 131–153.
- [3] X. Chen, X. Zhang, Z. Y. Zhao, J. H. Yang, V. Lantz, K. Q. Wang, "Hand Gesture Recognition Research Based on Surface EMG Sensors and 2D-accelerometers", *Proceedings of the 11th IEEE International Symposium on Wearable Computers*, Boston, MA, USA, 11-13 October 2007, pp. 11-14.
- [4] Y. J. Chang, S. F. Chen, J. D. Huang, "A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities", *Research in Developmental Disabilities*, Vol. 32, No. 6, 2011, pp. 2566–2570.
- [5] L. Omelina, B. Jansen, B. Bonnechère, S. Van Sint Jan, J. Cornelis, "Serious games for physical rehabilitation: designing highly configurable and adaptable games", *Proceedings of the 9th International Conference on Disability, Virtual Reality & Associated Technologies*, Laval, France, 10-12 September 2012, pp. 195–201.
- [6] D. Farina, N. Jiang, H. Rehbaum, A. Holobar, B. Graimann, H. Dietl, O. C. Aszmann, "The Extraction of Neural Information from the Surface EMG for the Control of Upper-Limb Prostheses: Emerging Avenues and Challenges", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 22, No. 4, 2014, pp. 797–809.
- [7] E. Scheme, K. Englehart, "Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use", *The Journal of Rehabilitation Research and Development*, Vol. 48, No. 6, 2011, 643–659.
- [8] S. Bai, J. Z. Kolter, V. Koltun, "An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling", *arXiv:1803.01271*, 2018.
- [9] P. Tsinganos, B. Cornelis, J. Cornelis, B. Jansen, A. Skodras, "Improved Gesture Recognition Based on sEMG Signals and TCN", *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Brighton, UK, 12-17 May 2019, pp. 1169–1173.
- [10] P. Tsinganos, B. Cornelis, J. Cornelis, B. Jansen, A. Skodras, "A Hilbert Curve Based Representation of sEMG Signals for Gesture Recognition", *Proceedings of the 26th International Conference on Systems, Signals and Image Processing*, Osijek, Croatia, 5-7 June 2019, pp. 201–206.
- [11] B. Hudgins, P. Parker, R. N. Scott, "A new strategy for multifunction myoelectric control", *IEEE Transactions on Biomedical Engineering*, Vol. 40, No. 1, 1993, pp. 82–94.
- [12] C. Castellini, A. E. Fiorilla, G. Sandini, "Multi-subject/daily-life activity EMG-based control of mechanical hands", *Journal of NeuroEngineering and Rehabilitation*, Vol. 6, No. 1, 2009, 41.
- [13] M. Atzori, et al, "Electromyography data for non-invasive naturally-controlled robotic hand prostheses", *Scientific Data*, 1, 140053, 2014.
- [14] A. Gijsberts, M. Atzori, C. Castellini, H. Muller, B. Caputo, "Movement Error Rate for Evaluation of Machine Learning Methods for sEMG-Based Hand Movement Classification", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 22, No. 4, 2014, pp. 735–744.
- [15] I. Kuzborskij, A. Gijsberts, B. Caputo, "On the challenge of classifying 52 hand movements from surface electromyography", *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, CA, USA, 28 August - 1 September 2012, pp. 4931–4937.
- [16] M. Atzori, et al, "Building the Ninapro database: A resource for the biorobotics community", *Proceedings of the IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechanics*, Rome, Italy, 24-27 June 2012, pp. 1258–1265.
- [17] K. H. Park, S. W. Lee, "Movement intention decoding based on deep learning for multiuser myoelectric interfaces", *Proceedings of the 4th International Winter Conference on Brain-Computer Interface*, Yongpyong, South Korea, 22-24 February 2016, pp. 1–2.
- [18] M. Atzori, M. Cognolato, H. Müller, "Deep Learning with Convolutional Neural Networks Applied to Electromyography Data: A Resource for the Classification of Movements for Prosthetic Hands", *Frontiers in Neurobotics*, Vol. 10, 2016, pp. 1–10.
- [19] P. Tsinganos, B. Cornelis, J. Cornelis, B. Jansen, A. Skodras, "Deep Learning in EMG-based Gesture Recognition", *Proceedings of the 5th International Conference on Physiological Computing Systems*, Seville, Spain, 19-21 September 2018, pp. 107–114.
- [20] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting", *Journal of Machine Learning Research*, Vol. 15, 2014, pp. 1929–1958.
- [21] W. Geng, Y. Du, W. Jin, W. Wei, Y. Hu, J. Li, "Gesture recognition by instantaneous surface EMG images", *Scientific Reports*, Vol. 6, No. 1, 36571, 2016.

- [22] W. Wei, Y. Wong, Y. Du, Y. Hu, M. Kankanhalli, W. Geng, "A multi-stream convolutional neural network for sEMG-based gesture recognition in muscle-computer interface", *Pattern Recognition Letters*, Vol. 119, 2019, pp.131–138.
- [23] N. Jiang, I. Vujaklija, H. Rehbaum, B. Graitmann, D. Farina, "Is Accurate Mapping of EMG Signals on Kinematics Needed for Precise Online Myoelectric Control?", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 22, No. 3, 2014, pp. 549–558.
- [24] S. Muceli, N. Jiang, D. Farina, "Extracting Signals Robust to Electrode Number and Shift for Online Simultaneous and Proportional Myoelectric Control by Factorization Algorithms", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 22, No. 3, 2014, pp. 623–633.
- [25] A. Stango, F. Negro, D. Farina, "Spatial Correlation of High Density EMG Signals Provides Features Robust to Electrode Number and Shift in Pattern Recognition for Myocontrol", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 23, No. 2, 2015, pp. 189–198.
- [26] Y. Du, W. Jin, W. Wei, Y. Hu, W. Geng, "Surface EMG-Based Inter-Session Gesture Recognition Enhanced by Deep Domain Adaptation", *Sensors*, Vol. 17, No. 3, 458, 2017.
- [27] Y. Li, N. Wang, J. Shi, J. Liu, X. Hou, "Revisiting Batch Normalization For Practical Domain Adaptation", arXiv:1603.04779, 2016.
- [28] U. Côté-Allard, et al, "Deep Learning for Electromyographic Hand Gesture Signal Classification Using Transfer Learning", arXiv:1801.07756, 2018.
- [29] P. Tsinganos, B. Cornelis, J. Cornelis, B. Jansen, A. Skodras, "Data Augmentation of Surface Electromyography for Hand Gesture Recognition," *Sensors*, Vol. 20, No. 17, 4892, 2020.
- [30] S. Dhahbi, W. Barhoumi, J. Kurek, B. Swiderski, M. Kruk, E. Zagrouba, "False-positive reduction in computer-aided mass detection using mammographic texture analysis and classification", *Computer Methods and Programs in Biomedicine*, Vol. 160, 2018, pp. 75–83.
- [31] J. Kurek, B. Swiderski, S. Osowski, M. Kruk, W. Barhoumi, "Deep learning versus classical neural approach to mammogram recognition", *Bulletin of the Polish Academy of Sciences: Technical Sciences*, Vol. 66, No. 6, 2018, pp.831–844.
- [32] T. Corcoran, R. Zamora-Resendiz, X. Liu, S. Crivelli, "A Spatial Mapping Algorithm with Applications in Deep Learning-Based Structure Classification", arXiv:1802.02532, 2018.
- [33] J. H. Bappy, C. Simons, L. Nataraj, B. Manjunath, A. K. Roy-Chowdhury, "Hybrid LSTM and Encoder-Decoder Architecture for Detection of Image Forgeries", *IEEE Transactions on Image Processing*, Vol. 28, No. 7, 2019, pp. 3286–3300.
- [34] B. Yin, M. Balvert, D. Zambrano, A. Schönthuth, S. Bohte, "An image representation based convolutional network for DNA classification", arXiv:1806.04931, 2018.
- [35] A. Thabet, H. Alwassel, B. Ghanem, "MortonNet: Self-Supervised Learning of Local Features in 3D Point Clouds", arXiv:1904.00230, 2019.
- [36] D. Jasrasaria, E.O. Pyzer-Knapp, D. Rappoport, A. Aspuru-Guzik, "Space-Filling Curves as a Novel Crystal Structure Representation for Machine Learning Models", arXiv:1608.05747, 2016.
- [37] C. Gotsman, M. Lindenbaum, "On the metric properties of discrete space-filling curves", *IEEE Transactions on Image Processing*, Vol. 5, No. 5, 1996, pp. 794–797.
- [38] B. Moon, H. V. Jagadish, C. Faloutsos, J. Saltz, "Analysis of the clustering properties of the Hilbert space-filling curve", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 13, No. 1, 2001, pp. 124–141.
- [39] D. Holzmüller, "Efficient Neighbor-Finding on Space-Filling Curves", arXiv:1710.06384, 2017.
- [40] P. Tsinganos, A. Skodras, B. Cornelis, B. Jansen, "Deep Learning in Gesture Recognition Based on sEMG Signals", *Learning Approaches in Signal Processing*, 1st Edition, 2018, p. 471.
- [41] T. T. Um, et al, "Data augmentation of wearable sensor data for parkinson's disease monitoring using convolutional neural networks", *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, Glasgow, UK, 13-17 November 2017, pp. 216–220.
- [42] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", arXiv:1409.1556, 2014.
- [43] G. Huang, Z. Liu, L. van der Maaten, K. Q. Weinberger, "Densely Connected Convolutional Networks", arXiv:1608.06993, 2016.
- [44] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size", arXiv:1602.07360, 2016.
- [45] C. Attig, N. Rauh T. Franke, J. F. Krems, "System Latency Guidelines Then and Now – Is Zero Latency Really Considered Necessary?", *Lecture Notes in Computer Science*, Springer, Vol. 10276, 2017, pp. 3–14.

Characterisation of Signal Amplitude-Frequency for Indoor Power Line Communication Channel in the 1 — 30 MHz Broadband Frequencies

Original Scientific Paper

Adedayo Oluyomi Ajibade

Department of Electrical/Electronic and Computer Engineering,
College of Engineering, Afe Babalola University
ajibadea@abuad.edu.ng

Ilesanmi Banjo Oluwafemi

Department of Electrical and Electronic Engineering,
Faculty of Engineering Ekiti State University,
ilesanmi.oluwafemi@eksu.edu.ng

Israel Esan Owolabi

Department of Electrical/Electronic and Computer Engineering,
College of Engineering,
Afe Babalola University
eowolabi2002@yahoo.com

Abstract – *The transmission of data signals over power lines is a very promising technique for delivering indoor broadband communication services. However, since power grids were originally designed for high-voltage low-frequency signal transmission, there is a frequency mismatch between the power grid and high-frequency data signals. This mismatch poses a challenge to deploying power lines as a communication channel. Although, studies and researches conducted in several countries have made transmission of data over power lines possible, the behaviour and properties of the power grid cannot be generalised. Hence, the need for in-depth experimental measurement and analysis on the suitability and capability of the Nigerian power grid for data transmission is crucial for proper characterising and modelling of the power line communication (PLC) channel. In this paper, we present experimental measurement results of the effects of frequency variations on the attenuation experienced by broadband high-speed data signals transmitted over the Nigerian indoor power line network.*

Keywords – *amplitude, attenuation, broadband frequency, channel, indoor PLC, signal*

1. INTRODUCTION

The ubiquitous nature of the power grid has attracted huge interest into its deployment as an alternative and efficient technique for delivering high-speed data signals to end-users around the world [1]. With the ever-increasing global demand of high-speed multimedia data services, the successful deployment of power lines as a communication channel is imperative [2], [3]. This will undoubtedly complement existing technologies currently deployed in many countries of the world. Although, power lines have been used for transmitting communication signals for several decades, its use has only been limited to transmission of low-rate control and monitoring signals by power supply companies. In

recent times, power line communication (PLC) technology has also been deployed in smart grid and for automatic meter reading technology [4], [5].

However, the power line network poses a harsh and unconducive environment as channel medium for the transmission of high-speed data signals due to the frequency and voltage mismatch between the two [6]. As a result, data signals undergo mild to severe attenuation as they transverse the power line channel from one point (node) to another. This attenuation is undesired in any communication system, and serves as the basis for power line communication (PLC) channel measurement, characterisation and modelling [6], [7].

Asides frequency and voltage mismatch, certain properties of the power line network contribute to signal attenuation on the power line communication (PLC) channel. Multipath propagation, noise, topology as well as node (load or no load) terminations play a major role in determining the extent of signal attenuation, and by extension, the efficacy of PLC channels for data transmission and delivery [7], [8].

While PLC channels have been characterized in varying capacities in different countries around the world, the characteristics and model(s) obtained in one country cannot be generalised for others [5], [9]. Furthermore, most of the existing literature on power line channel modelling focus on the deterministic approach, where the power line is analysed as a transmission line.

Some research works have been presented in literature on in-home PLC. In [10], characterisation of some urban and sub-urban in-home PLC channels in the 2 – 30 MHz frequency band in terms of the average channel gain (ACG) and the root mean squared delay-spread (RMS-DS) in United State of America was presented. While in [9] and [11], the results for the evaluation of in-home PLC channels – where the average channel gain (ACG), coherence bandwidth (CB) and the root mean squared delay-spread (RMS-DS) were used to evaluate the channel for France were presented. The results of the channel attenuation and noise for frequencies up to 30 MHz in Spain were presented in [12] and [10] where the ACG, RMS-DS and CB were presented. Also, analysis of the ACG, RMS-DS and CB for the upper frequency edge of 300 MHz was carried out in Spain in [11]. In [12] and [13], extensive measurement campaign and numerical analysis of the ACG, RMS-DS and CB parameters for the Brazilian in-home power line channel were presented. In this research work, the results of preliminary measurements on the Nigerian indoor PLC channel are presented. This is to provide a reference for subsequent characterization and modeling of indoor broadband PLC systems in Nigeria.

The remainder of this paper is organized as follows: Section II discusses the Nigerian indoor power line communication channel while section III describes the equipment set up for the experiment and measurement procedures. The results obtained from the experiment are analysed and discussed in section IV; and finally, conclusions are presented in section V.

1.1 Comparative Analysis of General Wired, Wireless and Power Line Communication Systems

Power line communication (PLC) is a form of wire communication, except that the wire medium used are not typical ones like Ethernet data cables or telephone coaxial cables. Rather, existing electric power cables are utilised for transmission of information signals [14], [15]. In the Nigerian scenario, wireless communication has been extensively deployed by mobile telecommu-

nication service vendor that have a very high presence in both urban and rural communities. Internet access is pretty much available as long as subscribers can afford internet subscription. On the other hand, communication over cables have not really been successful, as it is obtainable in developed countries [16]. This is largely due to unavailability of cable infrastructures that ought to have been put in place. Private establishments like institutions of learning, large corporations and few individuals appear to be the only group that can afford the high cost, and other modalities, required for setting up and maintaining these systems. As a result, hitherto, internet user traffic are tilted towards wireless technologies than the wired counterpart.

The potential benefits of deploying PLC systems in Nigeria include cost effectiveness i.e. the use of existing power cables; ease of network setup; huge bandwidth available at the frequency of operation; low latency and high reliability [15], [16]. All these contribute to making PLC technology an excellent candidate for meeting future broadband internet access demands in Nigeria [17].

1.2 Modulation Techniques for Power Line Communication Systems

For power line communication systems, digital modulation techniques, including frequency-shift keying (FSK), Phase-shift keying (PSK), and quadrature-amplitude modulation (QAM) have been deployed as modulation scheme for PLC systems [18]. However, due to the multipath characteristics of power line communication systems, a more suitable modulation technique is the orthogonal frequency-division multiplexing (OFDM) [19], [20].

Although, the separate subcarriers in OFDM overlap, they are orthogonal to one another; the orthogonality implies that the peak of one subcarrier occurs when the others are at zero. Intersymbol interference (ISI) is mitigated in OFDM by inserting a cyclic prefix (CP) (or guard interval) [21], [22]. The insertion of CP also permits proper demodulation of the symbol streams without spectrum overlap [22]. Asides mitigation of ISI in OFDM, the technique is very reliable in multipath propagation environments like the PLC network [23], [24]. It offers very high resistance to fading and can accommodate delay spread induced by multipath propagation [19], [25].

1.3 Potential Applications of Power Line Communication in Nigeria

The huge population of Nigeria, coupled with the continuous increase in the demand for internet and multimedia services have necessitated the need for complementary technologies and services that will ensure perennial internet access [16]. In about 20 years of mobile communication systems' existence in Nigeria, the number of active mobile internet subscribers has risen from just about a million to over 50 million. Pres-

ently, Nigeria ranks highest in Africa in internet accessibility, and among the top 20 in the world [16], [17]. Even with the deployment of fourth generation (4G) mobile networks, the existing networks are insufficient to meet prevailing demand. It is envisaged that in the next few years, demand for broadband internet services may exceed network providers' capacity.

Power line communication (PLC) system offers an alternative technology for providing reliable broadband internet access to Nigerian residential homes, offices and industrial environment. Although, only about 40 per cent of Nigerians have access to power - which are largely in urban areas and cities, government are putting infrastructures in place to make sure that at least 80% of its population have access to electricity within the next few years [16]. Asides electric power challenges, telephones and data cable connections are severely limited, even in urban regions and cities. When these power infrastructures are in place, PLC deployment will ensure that people in remote rural areas and urban centres will have access to power. However, even with access to power, broadband internet access in rural and remote areas are rare or non-existent, in most cases [16], [17]. Thus, as long as there is connection to the electric power grid, PLC systems can be deployed to provide reliable and affordable internet access over the power line [26]. In light of this, characterisation and modelling of the Nigerian PLC channel is important in preparation for the deployment of in-home PLC in Nigeria.

2. OVERVIEW OF THE NIGERIAN POWER LINE NETWORK

The availability and reach of electric power cables make it an alternative medium for transmission of data signals [27]. Since the power line infrastructures (power grid) are already in existence in virtually all countries of the world, the deployment cost for PLC systems is reduced drastically when compared to other communication technologies [27], [28]. While PLC channels display strong resilience against natural hazards, power lines are still capable of transmitting low-voltage communication signals even with the occurrence of faults that may render them incapable of transmitting high-voltage electric power signals [29]. However, PLC channels exhibit certain propagation characteristics that affect signal-carrying capability [29], [30]. These characteristics in relation to the Nigerian indoor power line networks are discussed in this section.

The Nigerian low-voltage indoor power line network has a last-mile alternating current (AC) voltage of 240 V, 60 Hz (one-phase) and 415 V, 60 Hz (three-phase). This experiment was restricted to single-phase measurements only. As shown in Fig. 1, indoor power line networks usually terminate in socket outlets that are connected to derivation boxes (DBs) in either bus or star topology [31]. All DBs in a building are then connected to a service panel that feeds energy from the service provider into the building. Joints in the power cables

and connection boxes, as well as series connection of cables with different characteristic impedances result in reflections (or echoes) on the PLC channel [31], [32]. Also, socket outlets are terminated in load (appliance) impedances of varying values or are open-circuited. Load impedances on the channel affects the performance by altering the channel frequency response and input impedance, and also contribute to the level of background noise [31], [33].

Thus, due of the contributions of connected impedances to background noise, additive white Gaussian noise (AWGN) is not sufficient in modelling PLC noise [31], [34]. Multipath propagation due to reflections on the power line, varying load impedance terminations and peculiar noise behaviour of the PLC channel all contribute to the attenuation of signals as they transverse the power line channel [29], [35], [36].

The experiment and measurements in subsequent sections show the variations in the level of attenuation experienced by signals in the indoor broadband frequencies of 1 – 30 MHz on a PLC channel.

3. EXPERIMENTAL METHOD AND MEASUREMENT PROCEDURES

The equipment set up for the measurement is shown in Fig. 2 and Fig. 3. The experiment was carried out in two different laboratory rooms in the Electrical/Electronic and Computer Engineering Department, College of Engineering of Afe Babalola University, Ado Ekiti, Nigeria.

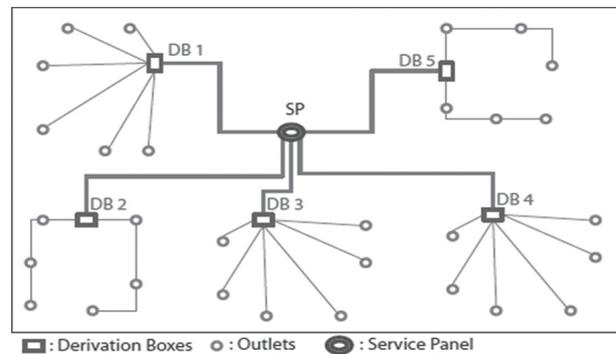


Fig. 1. Indoor power line distribution [19]

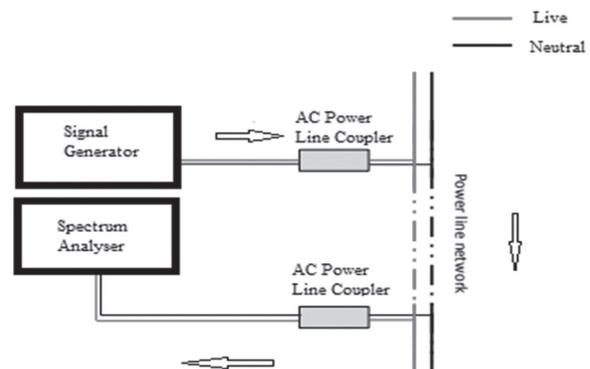


Fig. 2. Block representation of the PLC channel measurement setup

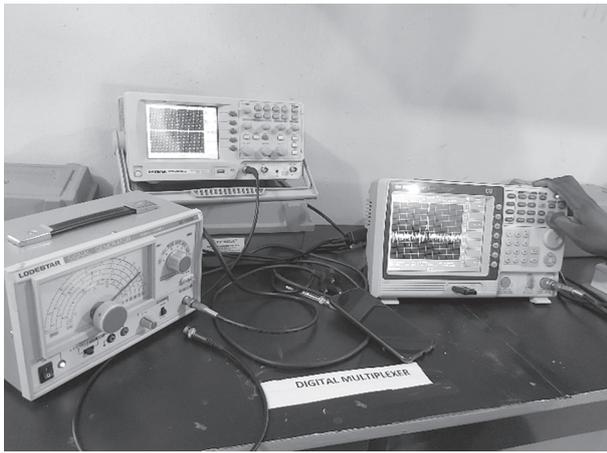


Fig. 3. Pictorial diagram of the PLC channel measurement setup

As shown, the set up consists of a function generator, a spectrum analyser, a pair of AC power line couplers and an indoor AC power line that links two socket outlets. Low pass filter AC coupling circuits were used to connect both the function generator and spectrum analyser to the power line. Besides isolating the measuring equipment from the mains high-voltage power network, these couplers also function to protect the measuring equipment from the 60 Hz AC frequency, and provide points of entry and exit for the low-voltage high-frequency signals respectively, into and from the power network.

The experiment was conducted over a fixed PLC channel cable length of about 5 m between two indoor socket outlets. The spectrum analyser measured and displayed the attenuation (in decibels) experienced by signals traversing the power line at particular frequencies. The waveforms were then saved as trace files on the spectrum analyser; and were processed and analysed using MATLAB and Excel software.

Three loads connected for this experiment include: a 20 W energy-saving bulb (a capacitive load); an 8 W mobile phone charger and a 65 W laptop charger (inductive loads). Signals at specific frequencies were injected in steps from the narrowband range of about 1.5 MHz up to the broadband range of 30 MHz into the system – under both no-load and load conditions. Also, the attenuation-frequency waveform for each level of signal frequency was obtained. The bandwidth and sweep for each level of injected signal were 3 MHz and 0.24 seconds respectively, with the frequency of interest at the centre. The injected signals were increased in steps of 5 MHz up to 30 MHz.

4. RESULTS AND DISCUSSION

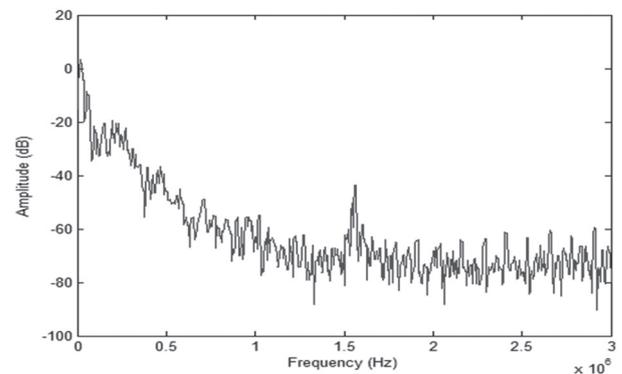
In this section, the data and waveforms obtained from the broadband frequency measurement up to 30 MHz are presented, and their variations with the corresponding signal power are discussed. Also, the results of the statistical computation of the average channel

attenuation are presented, and comparison with results available from existing literature are made.

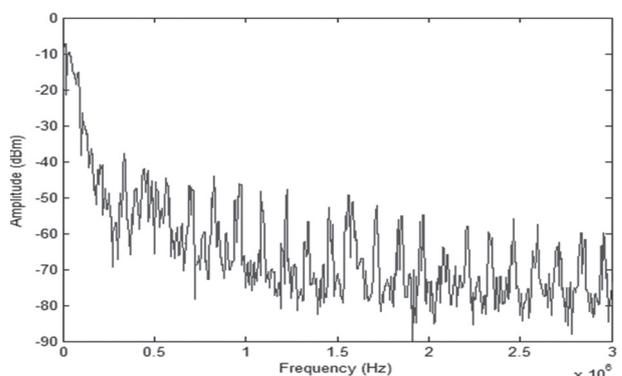
4.1 Received Signal Amplitude versus frequency for No-load and Load Conditions

Fig 4 shows the Amplitude-Frequency waveform at frequency up to 1.5 MHz. It can be observed that at the lower frequencies up to 1.5 MHz, the amplitude of the received signal dropped steadily to very low values of -42 dBm and -52dBm for both no-load and load conditions respectively. This implies that the severity of the attenuating effects of the PLC channel on transmitted signal at frequencies within this range rapidly increased. From 500 Hz, the signal amplitude declined in a fairly stable manner until 1.5 MHz, when the amplitude improved to about -39 dB and -46 dBm for no-load and load conditions respectively.

The Amplitude-Frequency waveform at 5 MHz is shown in Fig. 5. It can be observed that at 5 MHz, the amplitude of the received signal improved further to about -24 dB for no-load and -25 dBm for load conditions. This implies a much less attenuation experienced by the signal at 5 MHz. Also, the attenuation values for both conditions are almost equal at this frequency i.e. the PLC channel effects on the transmitted signal was load-independent.

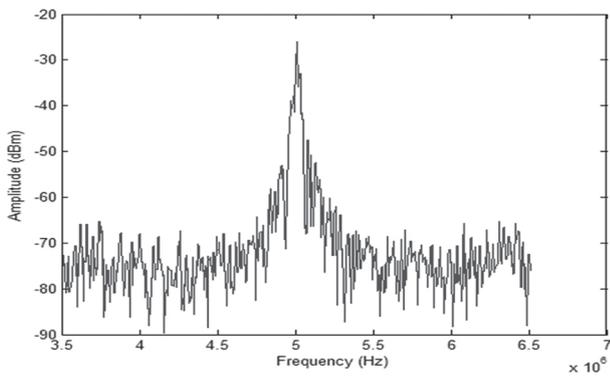


(a) no-load condition

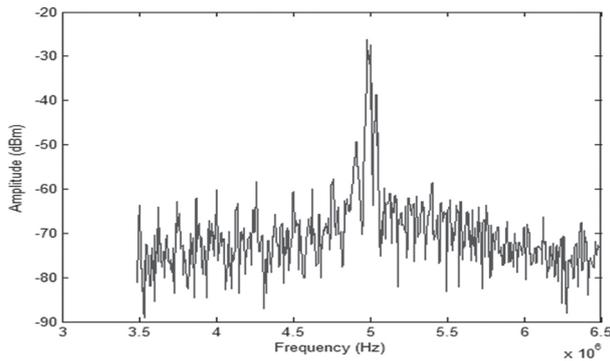


(b) load condition

Fig. 4. Amplitude-frequency waveform at 1.5 MHz

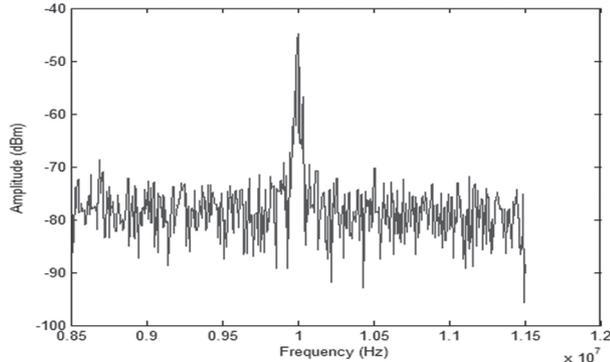


(a) no-load condition

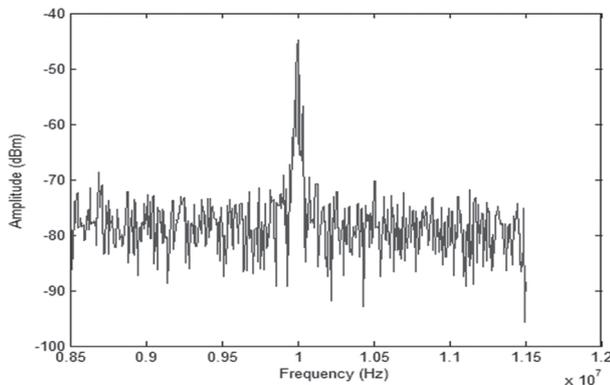


(b) load condition

Fig. 5. Amplitude-frequency waveform at 5 MHz

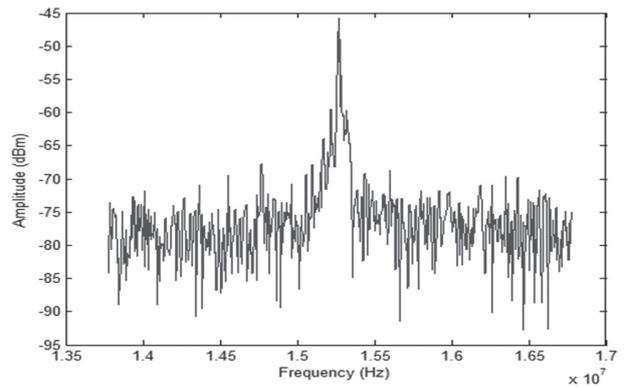


(a) no-load condition

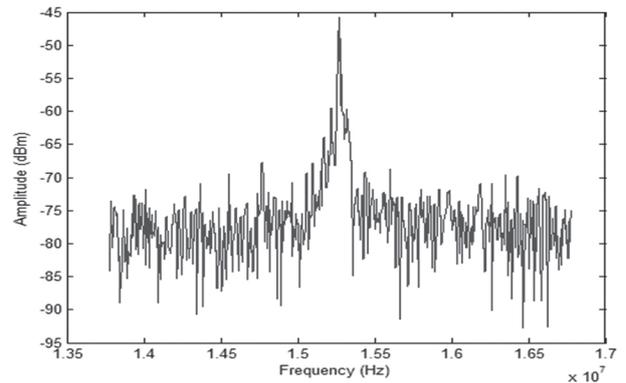


(b) load condition

Fig. 6. Amplitude-frequency waveform at 10 MHz



(a) no-load condition



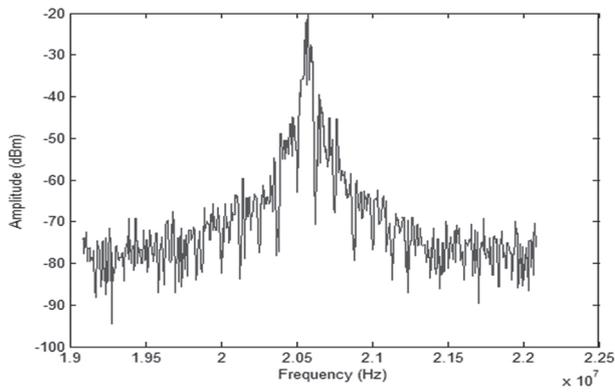
(b) load condition

Fig. 7. Amplitude-frequency waveform at 15 MHz

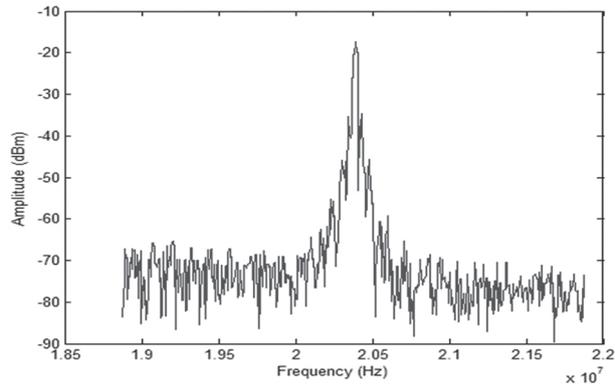
Fig. 6 shows the waveform at 10 MHz, where it can be observed that the received signal amplitude further dropped to -43 dBm and -42 dBm for no-load and load conditions respectively, as can be observed from Fig. 6. The received signal amplitude for both load conditions are approximately the same i.e. the load conditions had negligible effects on the PLC channel attenuation at this frequency as well.

Similar to the results obtained for 5 MHz and 10 MHz transmitted signal, the attenuation experienced by the signal over the PLC channel at 15 MHz was observed to be quite comparable for the no-load and load conditions, as shown in Fig. 7, i.e. -46 dBm for the former and -43 dBm for the latter. Furthermore, it can be observed that the signal attenuation increased steadily for signals at 5 MHz through 10 MHz to 15 MHz for both load conditions.

At 20 MHz signal frequency, the attenuation experienced by the signal traversing the PLC channel did not deviate significantly from each other. However, at this frequency, the signal amplitude greatly improved to -18 dBm and -20 dBm for no-load and load conditions respectively, as seen in Fig. 8. This indicates that the PLC channel had the least attenuating effect so far on the transmitted signal, as the signal amplitudes were highest for all signal frequencies up to 20 MHz.

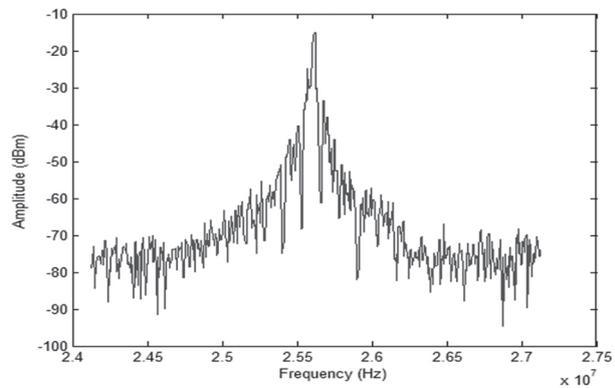


(a) no-load condition

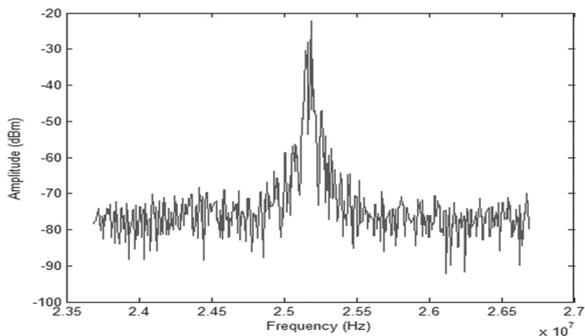


(b) load condition

Fig. 8. Amplitude-frequency waveform at 20 MHz

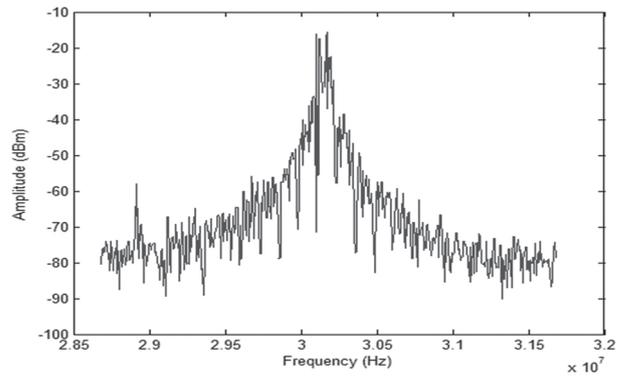


(a) no-load condition

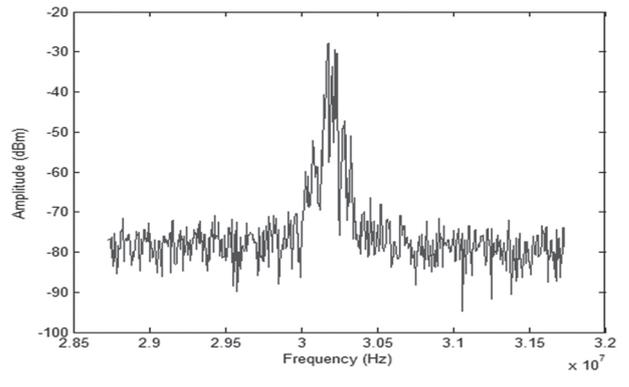


(b) load condition

Fig. 9. Amplitude-frequency waveform at 25 MHz



(a) no-load condition



(b) load condition

Fig. 10. Amplitude-frequency waveform at 30 MHz

As can be observed from Fig. 9, at 25 MHz, the amplitude of the received signal for the PLC channel on load was much higher than without load. With amplitude of -14 dBm on no-load and -22 dBm on load, the load conditions on the PLC channel had a huge attenuating impact on the signals traversing the channel at this frequency.

Lastly, at 30 MHz signal frequency, Fig.10 shows that the disparity in the amplitude of the received signal has further widened from -16 dBm for the no-load PLC channel condition to -28 dBm for the loaded condition. This means that the load on the PLC channel induced greater attenuation for signals at 30 MHz frequency than at all previous sampled frequencies.

4.2 Variations in the received signal power with frequency for no-load and load conditions

Table 1 shows the variations in the received signal power with frequency for the no-load and on-load PLC channel. From the table, for both load and no-load conditions, it can be observed that the received signal power is very low from narrowband range up to about 5 MHz signal frequency i.e. the point the signal power started increasing. The implication is that the channel attenuation is more profound at this low frequency range. But, at 10 MHz and 15 MHz signal frequencies,

the signal attenuation became high again, resulting in low signal power within this range.

This conforms to the communication theory that signal attenuation increases with frequency in a non-ideal conductor. Finally, from 20 MHz up to 30 MHz frequencies, the received signal power improved significantly for both no-load and load conditions.

Table 1 Variations in the received signal power with frequency for PLC channel on no-load and load conditions

Frequency (MHz)	Signal power (mW) (no-load)	Signal power (mW) (load)
0.5	0.00006	0.00001
1.5	0.00013	0.00003
5	0.00398	0.00316
10	0.00005	0.00006
15	0.00003	0.00005
20	0.01585	0.01000
25	0.03981	0.00631
30	0.02519	0.00159

4.3 Statistical measurement of the average attenuation on the PLC channel

The average channel attenuation (ACA) is a well-established parameter used in characterizing indoor PLC channels. The average channel attenuation is the negative of the average channel gain (ACG), and is expressed mathematically as [37]:

$$ACA = -ACG = -10 \log_{10} \left(\frac{1}{N} \sum_{k=0}^{N-1} |H[k]|^2 \right) \quad (1)$$

Where N is the number of subcarriers and $H[k]$ is the k th coefficient of the discrete Fourier transform of the linear time-invariant PLC channel impulse response.

For our PLC channel setup, using Eq. (1), the ACA was determined by considering channel frequency response (CFR) estimates over 1 – 30 MHz frequency range. Several CFR estimates were taken over varying length of a.c. socket outlet pairs in the laboratory rooms, each with its own range of connected equipment and appliances. The overall values for the maximum, minimum and mean of the average channel attenuation (ACA) in dB are presented, as shown in Table 2.

The table also compares the ACA values with those that exist in literature from other countries [13], [37].

Table 2. Comparison of average channel attenuation (ACA) obtained for the Nigerian 1 – 30 MHz indoor broadband PLC channel with other countries

Statistical Parameters	ACA (dB) Nigeria	ACA (dB) Brazil	ACA (dB) U. S.	ACA (dB) Spain
Maximum	76.86	51.09	68.12	~70
Minimum	16.15	9.18	19.70	~10
Mean	43.5	23.28	48	~30

From table 2, it can be observed that all the three statistical parameters – maximum, minimum and mean values of the ACA for Nigerian PLC channel are higher than what is obtained in Brazil and Spain. However, the maximum ACA value for Nigerian PLC is higher than that of U.S.A. while the mean and minimum values are less than the U. S. A. Thus, the Nigerian PLC channel has the highest maximum ACA (dB) value among all four countries considered in the frequency band of interest (i.e. 1 – 30 MHz). Also, the minimum and mean ACA (dB) are higher than that of Spain and Brazil. The high value obtained for maximum ACA may be attributed to the nature and peculiarities of the power grid besides from other ‘noisy’ equipment that may be connected to the larger power network while measurements were being taken. By extending the measurements campaign further, in order to obtain a wider data set, the ACA values may change. The final values obtained will be very crucial to characterising and modelling the Nigerian PLC channel.

5. CONCLUSIONS

A lot of factors influence the behavior of indoor power line channels. In this work, we have presented the results of experimental measurements on the Nigerian indoor power network, and considered the effects of channel attenuation on signals transmitted across the channel under load and no-load PLC network scenarios. Some conclusions are derived from this work. The narrowband frequencies up to 1 kHz are capable of transmitting high-speed data signals for both conditions of load over the power line channel without significant attenuation. Likewise, higher broadband frequencies from 20 to 30 MHz are shown to offer better immunity to the attenuating effects of the PLC channel on transmitted signals, thereby offering better capacity to transmit information signals. Further measurements at 1 – 30 MHz broadband frequencies with real-time traffic data will allow for characteristics and subsequent modelling of the Nigerian PLC channel for high-speed data transmission.

6. REFERENCES

- [1] L. T. Berger, A. Schwager, J. J. Escudero-Garzás, "Power Line Communications for Smart Grid Applications", *Journal of Electrical and Computer Engineering*, Vol. 2013, 16 pages.
- [2] Z. Mingyue, "Measurements and Channel Characteristics of LV Power Line Communications Networks in China", *Proceedings of the IEEE International Symposium on Power Line Communications and Its Applications*, Orlando, FL, USA, 26-29 March 2006, pp. 212–216.
- [3] W. Zhu, X. Zhu, E. Lim, Y. Huang, "State-of-art power line communications channel modelling", *Procedia Computer Science*, Vol. 17, 2013, pp. 563–570.
- [4] S. Guzelgoz, "Characterizing Wireless and Power-line Communication Channels with Applications to Smart Grid Networks", University of South Florida, 2011, Ph.D. Thesis.
- [5] A. M. Tonello, "PLC for the Smart Grid : State-of-the-Art and Challenges", *Proceedings of the 4th international Conference on Communications, Mobility and Computing*, Guilin, China, 21-23 May 2012, pp. 2–5.
- [6] M. Mosalaosi, "Power Line Communication (PLC) Channel Measurements and Charaterization", University of Kwazulu-Natal, 2014, M.Sc. Thesis.
- [7] F. Zwane, "Power Line Communication Channel Modelling", University of Kwazulu-Natal, 2014, M.Sc. Thesis.
- [8] A. M. Tonello, A. Pittolo, M. Girotto, "Power line communications: Understanding the channel for physical layer evolution based on filter bank modulation", *IEICE Transactions on Communications*, Vol. E97-B, No. 8, 2014, pp. 1494–1503.
- [9] A. M. Tonello, S. Member, F. Versolatto, S. Member, A. Pittolo, "In-Home Power Line Communication Channel : Statistical Characterization", *IEEE Transactions on Communications*, Vol. 62, No. 6, 2014, pp. 2096–2106.
- [10] F. J. Canete Corripo, J. A. Cortes, L. Diez, J. T. Entrambasaguas, "A Channel Model Proposal for Indoor Power Line Communications", *IEEE Communications Magazine*, Vol. 49, No. December, 2011, pp. 166–174.
- [11] A. M. Tonello, F. Versolatto, B. Béjar, S. Zazo, "A fitting algorithm for random modeling the PLC channel", *IEEE Transactions on Power Delivery*, Vol. 27, No.3, 2012, pp. 1477–1484.
- [12] T. R. Oliveira, C. B. Zeller, S. L. Netto, M. V. Ribeiro, "Statistical modeling of the average channel gain and delay spread in in-home PLC channels", *Proceedings of the IEEE International Symposium on Power Line Communications and Its Applications*, Austin, TX, USA, 29 March-1 April 2015, pp. 184–188.
- [13] T. Oliveira, A. Picorone, C. Zeller, S. Netto, M. Ribeiro, "Statistical Modeling of Brazilian In-Home PLC Channel Features", *Journal of Communication and Information Systems*, Vol. 34, No. 1, 2019, pp. 154–168.
- [14] S. Shukla, et al, "Comparison of Wireless Network over Wired Network", *International Journal of Research Granthaalayah*, vol. 5, 2017, pp. 14–20.
- [15] A. O. Ajibade, I. B. Oluwafemi, A. O. Ojo, K. A. Adeniji, "Theoretical Analysis of Transmission Parameters and Interference Issues in Power Line Communication Systems", *ABUAD Journal of Engineering Research and Development*, Vol. 1, No. 1, 2017, pp. 95–99.
- [16] O. J. Femi-Jemilohun, "Broadband Internet Access Services through Power Line Communication to the Rural Dwellers in Nigeria", *International Journal of Engineering Research & Tehnology*, Vol. 8, No. 4, 2019, pp. 112–115.
- [17] L. K. Matthew, F. Olowononi, J. S. Soja, "Power Line Communications: A Platform for Sustainable Development", *Proceedings of the 3rd International Conference On African Development Issues*, Ota, Nigeria, 9-11 May 2016, pp. 45–61.
- [18] A. Recioui, "Application of multiple input multiple output power line communication (MIMO-PLC) to power systems", *Acta Physica Polonica A*, Vol. 132, No. 3, 2017, pp. 496–499.
- [19] W. Zhang, "Design of Power Line Carrier Communication System Based on Orthogonal Frequency Division Multiplexing", *Proceedings of the 5th International Conference on Advanced Materials and Computer Science*, Qingdao, China, 26-27 March 2016, pp. 784–787.

- [20] T. D. Farkas, T. Király, T. Pardy, T. Rang, G. Rang, "Application of power line communication technology in street lighting control", *International Journal of Design & Nature and Ecodynamics*, Vol. 13, No. 2, 2018, pp. 176–186.
- [21] F. Pavliodou, N. Sad, "Power line communications based on OFDM 1", *Viatacom Ltd.*, Vol. 2, No. 1, 2013, pp. 1–4.
- [22] E. M. Meza, J. B. Aspiazu, "Design of a Power Line Communications Transceiver Based on OFDM," *Procedia Technology*, Vol. 17, 2014, pp. 107–113.
- [23] N. A. M. Mohammed, "Design of Power Line Communication", *University of Khartoum, India*, 2017, B.Sc. Thesis.
- [24] N. Andreadou, F. N. Pavlidou, "Modeling the noise on the OFDM power-line communications system", *IEEE Transactions on Power Delivery*, Vol. 25, No. 1, 2010, pp. 150–157.
- [25] A. Pandharipande, "Principles of OFDM", *IEEE Potentials*, Vol. 21, No. 2, 2002, pp. 16–19.
- [26] S. Mehfuz, "Performance Analysis and Modeling of Broadband Power line Communications", *International Journal of Advanced Research in Electronics and Communication Engineering*, Vol. 1, No. 2, 2012, pp. 22–27.
- [27] L. T. Berger, A. Schwager, P. Pagani, D. M. Schneider, "MIMO Power Line Communications", *IEEE Communications Surveys & Tutorials*, Vol. 17, No. 1, 2014, pp. 106–124.
- [28] N. Sagar, "Powerline Communications Systems: Overview and Analysis", *New Brunswick, New Jersey*, 2011, M.Sc. Thesis
- [29] N. Weling, N. Nazari, "Statistical evaluation of 55 million PLC channel and topology measurements by more than 75,000 end-users", *Proceedings of the IEEE International Symposium on Power Line Communications and Its Applications*, Udine, Italy, 3-6 April 2011, pp. 237–242.
- [30] M.A. Sonmez, M. Bagriyanik, "Statistical model study for Narrowband Power Line Communication Noises", *Proceedings of the 8th International Conference on Electrical and Electronics Engineering*, Bursa, Turkey, 28-30 November 2013, pp. 539–543.
- [31] G. Marrocco, D. Statovci, S. Trautmann, L. A. Gmbh, A. Villach, "A PLC Broadband Channel Simulator for Indoor Communications", *Proceedings of the IEEE 17th International Symposium on Power Line Communications and Its Applications*, Johannesburg, South Africa, 24-27 March 2013, pp. 321–326.
- [32] A. M. Nyete, T. J. O. Afullo, I. E. Davidson, "Power line noise measurements and statistical modelling in the time domain", *Proceedings of PIERS*, Prague, Czech Republic, July 6-9 2015, pp. 1569–1574.
- [33] S. Khan, A. Salami, W. A. Lawal, Z. Alam, S. A. Haameed, M. J. Salami, "Characterization of Indoor Power lines As Data Communication Channels Experimental Details and Results", *International Journal of Electronics and Communication Engineering*, Vol. 2, No. 10, 2008, pp. 624–629.
- [34] P. Mlynek, J. Misurec, M. Koutny, "Noise modeling for power line communication model", *Proceedings of the 35th International Conference on Telecommunications and Signal Processing*, Prague, Czech Republic, 3-4 July 2012, pp. 282–286.
- [35] F. Aalamifar, A. Schlogl, D. Harris, L. Lampe, "Modelling power line communication using network simulator-3", *Proceedings of the IEEE Global Communications Conference*, Atlanta, GA, USA, 9-13 December 2013, pp. 2969–2974.
- [36] P. Mlynek, M. Koutny, J. Misurec, "Multipath channel model of power lines", *Elektrorevue*, Vol. 1, No. 2, 2010, pp. 48–53.
- [37] T. R. Oliveira, A. A. M. Picorone, S. L. Netto, M. V. Ribeiro, "Characterization of Brazilian in-home power line channels for data communication", *Electric Power Systems Research*, Vol. 150, 2017, pp. 188–197.

Steady-State Analysis Of Single Phase IPM Motors By D-Q Harmonic Balance Method

Original Scientific Paper

Asogwa Jude. C.

Energy Commission of Nigeria,
Abuja, Nigeria
jasogwa@yahoo.com

Obe Emeka. S.

University of Nigeria, Nsukka
Faculty of Engineering, Department of Electrical
Engineering
Enugu State, Nigeria
simon.obe@unn.edu.ng

Nnadi Damian. B.

University of Nigeria, Nsukka
Faculty of Engineering, Department of Electrical
Engineering
Enugu State, Nigeria
damian.nnadi@unn.edu.ng

Oti Stephen. E.*

University of Nigeria, Nsukka
Faculty of Engineering, Department of Electrical
Engineering
Enugu State, Nigeria
stephen.oti@unn.edu.ng
*Corresponding Author

Abstract – A concise steady-state analysis of a single-phase line-start permanent magnet (SPLSPM) machine is conducted from a developed d - q model using the d - q harmonic balance technique. The d - q model was developed in the rotor reference frame from a phase variable model of the machine. SPLSPM whose performance indices were characterized by high torque ripples has detailed analysis done mostly in computer simulations quite unlike the three-phase types. The main cause is not far-fetched, it was due to the nonexistence of a precise mathematical model in the d - q rotor frame of the motor due to the unbalanced field winding, the rotor saliency, and the presence of the capacitor in the auxiliary windings. Even after the model has been developed, the simple traditional procedure of setting all time-varying components to zero for steady-state analysis fails because the rotor position dependence on the inductance expressions could not be eliminated. The d - q harmonic-balance technique was then applied. The analysis and simulation were carried out using MatLab/Simulink software. An important feature of the harmonic balance technique was that it decoupled all equations to simple sine waveforms in a style that resembled the Fourier series. Results yield torque pulsation, current, and load characteristics in the steady-state.

Keywords – Torque, Capacitor, Permanent magnet, Transformation, Transient

1. INTRODUCTION

The high-efficiency rating of an interior permanent machine (IPM) among other single-phase motors makes it have the potential of replacing other classes of single-phase machines for both household and industrial applications. In the range of a few hundred watts, single-phase machines were customarily employed in the control, and in small energy drives, robotics and other associated usages which needed meticulous and high-power factor. Even though IPM synchronous motor is comparatively straightforward in building, it is rather very tough in analysis relative to its counterpart, 3- Φ machines due to asymmetrical nature of the motor as a result of unevenness in the field winding arrangement, irregular air-gap length, the presence of start and run capacitors. For the mere fact that its per-

formance index is characterized by high ripple torque, it tends to complicate further its analysis [1, 2]. In the process of carrying out a closed examination of the steady-state of IPM synchronous motor, it is quite relevant to be able to predict adequately the machine performance characteristics under a wide range of loading conditions [3]. Adequate studies were therefore performed for dynamic analysis. At present, there is not much literature available on this area and aspect of study so that an elaborate comparison would be made. Although [4,5] worked on Interior Permanent Motor, their concern was on other issues away from the d - q harmonic method. For instance [5] was interested in investigating the steady-state and dynamic response of an interior permanent magnet (IPM) synchronous machine drive to a single-phase open-circuit fault. In the past, Miller [6] developed a unified tactic in the analysis

of single-phase (1- Φ) machines conditioned on Fortescu's symmetrical component method which gave room for the application of intuitive reasoning. In like manner, [7] looked at the deployment of the technique in rotor current evaluation in respect of single-phase synchronous reluctance motor. However, this work attempts to present a detailed analytical model that will adequately predict single-phase machine IPM motor characteristics at a wide range of operating condition.

2. THE MATERIALS AND METHODOLOGY

Figure 1 shows the IPM machine made up of main and supplementary windings that were 900 electrical out of phase with each other. The capacitor was connected to auxiliary (supplementary) winding in series, in this way an additional phase was created from the 1- Φ supply. That connection made it possible for the machine to start directly online. The centrifugal switch as connected was to disconnect the start capacitor as soon as the motor has attained the required speed; the series capacitor ensured that the motor maintains a moderately high running torque, power factor with fewer torque pulsations [8]. By estimating the capacitance value of such capacitor used and applying such values for all the machine loading states, the optimum performance of all such machines is attained at all-times. For this analysis, the rated value of starting and running capacitor for the motor employed was 50 μF and 15 μF respectively. For the transition to synchronism, a précised switching speed be estimated such that the ripples produced at switching be minimal.

2.1 The Model of the Machine

Generally, as opined by [9], all single-phase machine transformations in the d-q axis were rotor position-dependent since the field windings were not balanced in the company of capacitor in the supplementary winding and the air gap not constant. The researchers in [10], thought that time-varying constraints cannot be completely removed by any reference frame transformations, however, they derived the d-q axis transformation for the capacitor voltage equation of the machine.

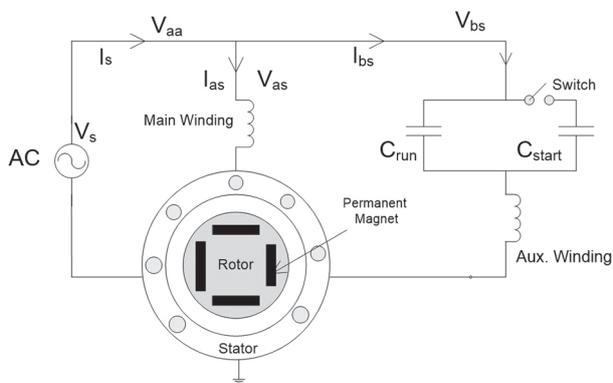


Fig.1. The diagram of the machine field winding and its rotor structure

$$V_{qs} = r_{aa}I_{qs} + r_{ab}I_{ds} + \omega_r \lambda_{ds} + p\lambda_{qs} + V_{qc} \quad (1)$$

$$V_{ds} = r_{bb}I_{ds} + r_{ab}I_{qs} - \omega_r \lambda_{qs} + p\lambda_{ds} + V_{dc} \quad (2)$$

$$V_{qr} = r_{qr}I_{qr} + p\lambda_{qr} \quad (3)$$

$$V_{dr} = r_{dr}I_{dr} + p\lambda_{dr} \quad (4)$$

$$pV_{qc} = -\omega_r V_{dc} + \frac{1}{2C_d} (I_{qs} - I_{ds} \sin 2\theta_r - I_{qs} \cos 2\theta_r) \quad (5)$$

$$pV_{dc} = \omega_r V_{qc} + \frac{1}{2C_d} (I_{ds} - I_{qs} \sin 2\theta_r + I_{ds} \cos 2\theta_r) \quad (6)$$

where,

$$r_{aa} = \frac{1}{2}(r_{as} + r_{bs}) + \frac{1}{2}(r_{as} - r_{bs}) \cos 2\theta_r \quad (7)$$

$$r_{bb} = \frac{1}{2}(r_{as} + r_{bs}) - \frac{1}{2}(r_{as} - r_{bs}) \cos 2\theta_r \quad (8)$$

$$r_{ab} = \frac{1}{2}(r_{as} - r_{bs}) \sin 2\theta_r \quad (9)$$

The under-listed equations are the flux linkage equations:

$$\lambda_{qs} = L_{lqs}I_{qs} + (I_{qs} + I_{qr})L_{mq} \quad (10)$$

$$\lambda_{ds} = L_{lds}I_{ds} + (I_{ds} + I_{dr})L_{md} + \lambda_m \quad (11)$$

$$\lambda_{qr} = L_{lqr}I_{qr} + (I_{qs} + I_{qr})L_{mq} \quad (12)$$

$$\lambda_{dr} = L_{ldr}I_{dr} + (I_{ds} + I_{dr})L_{md} + \lambda_m \quad (13)$$

2.2 The d-q Axis Modeling with Rotor Permanent Magnet

A rotor coil can be used to model a permanent magnet using a dc excitation current I_{rm} , such that $I_{rm} = E_{ro} / \omega_r L_{md}$, where E_{ro} is the excitation voltage due to magnet in the rotor since the product of magnetic flux linkage and the machine rotor speed, $\omega_r \lambda_m$ have the dimension of the excitation voltage. The permanent magnet gives a constant flux λ_m , where $\lambda_m = I_{rm} L_{md}$. Flux linkage in the field windings of the 1- Φ machine due to the permanent magnet in the rotor and which rotates in an unbalanced condition is modeled as [11]:

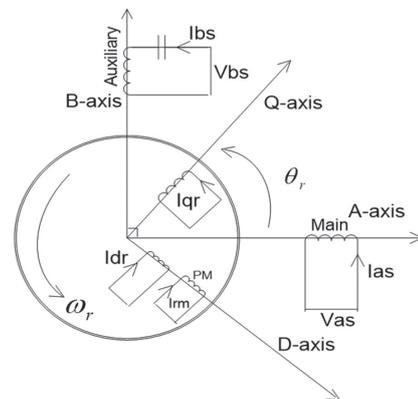


Fig. 2. Modeling of rotating IPM motor in the q-d axis reference frame

The q-d axis currents of the magnet torque can be derived once the IPM machine operates as a generator and the field winding is bridged, the established torque then became the braking torque produced by the presence of rotor magnets. Because of this, equations (14) and (15) were substituted in the q-d axis voltage equations (1) and (2) while setting the voltages to zero as the machine operates at a constant speed. The d-q axis magnet flux linkage equations, therefore, emerge as stated in equations (16) and (17) below [12].

$$\lambda_{qm} = L_{qs}i_{qm} \quad (14)$$

$$\lambda_{dm} = L_{ds}i_{dm} + \lambda_m \quad (15)$$

$$0 = r_{qs}i_{qm} + r_{qd}i_{dm} + \omega_r L_{ds}i_{dm} + \omega_r \lambda_m \quad (16)$$

$$0 = r_{ds}i_{dm} + r_{dq}i_{qm} - \omega_r L_{qs}i_{qm} \quad (17)$$

Solving equations (16) and (17) simultaneously led to the expressions for i_{qm} and i_{dm} in equations (18) and (19). These were useful for calculating the magnet torque.

$$i_{qm} = \frac{-\omega_r r_{ds} \lambda_m}{r_{ds} r_{qs} - (r_{qd} + \omega_r L_{ds})(r_{dq} - \omega_r L_{qs})} \quad (19)$$

$$i_{dm} = \frac{\omega_r \lambda_m (r_{dq} - \omega_r L_{qs})}{r_{ds} r_{qs} - (r_{qd} + \omega_r L_{ds})(r_{dq} - \omega_r L_{qs})} \quad (20)$$

2.3 The Machine Dynamic Torque Equations

At synchronous speed, the electromagnetic torque expression of the single-phase motor with a permanent magnet embedded in the rotor can therefore be expressed into two aspects of cage torque and magnet torque as shown below.

$$T_{em} = T_c + T_m = \frac{p}{4} (\lambda_{dr} i_{qr} - \lambda_{qr} i_{dr}) + \frac{p}{4} (\lambda_{dm} i_{qm} - \lambda_{qm} i_{dm}) \quad (21)$$

$$T_{em} = \frac{p}{4} (\lambda_{ds} I_{qs} - \lambda_{qs} I_{ds}) \quad (22)$$

$$T_c = \frac{p}{4} (\lambda_{dr} i_{qr} - \lambda_{qr} i_{dr}) \quad (23)$$

$$T_m = \frac{p}{4} (\lambda_{dm} i_{qm} - \lambda_{qm} i_{dm}) \quad (24)$$

The above-written equations (1 - 9, 10 - 13, and 20 - 25) enable researchers to carry out the fundamental frequency analysis [13] by the use of computer simulations of the machine to obtain the starting transients and steady-state waveforms. The cage and magnet torque of the machine during its asynchronous operation is also deduced in equation (20). The mechanical equations of such a machine (relating the rotor speed with the established electromagnetic torque) were described by equations (24) and (25).

$$J \left(\frac{2}{p} \right) p \omega_r = T_{em} - T_L \quad (24)$$

$$\theta_r = \frac{1}{p} \int_0^t \omega_r dt \quad (25)$$

3. APPLICATION OF HARMONIC BALANCE TECHNIQUES TO IPM MACHINE ANALYSIS

Since the stator winding is not well-adjusted and the air gap is not constant, considering the rotor saliency and the high torque ripples that characterize the performance indices of IPM machines, it has been noted earlier that the time-varying constraints of the motor cannot be removed completely by any reference frame transformation [3, 6]. It is also unequivocally tough to obtain the numerous current, flux linkage, and torque make-ups and their various sizes (the reason being that they were hooked on to the machine constraints and the capacitor values) without the application of phasor analysis. Those considerations as illustrated informed our decisions to develop an analytical model that will be able to address the presence of the third harmonic currents in the stator windings as a result of the asymmetrical nature of single-phase machines [13,11]. If phasor analysis was applied to the time fluctuating differential equation, the harmonic balance automatically decouples the state variables (which was assumed to be indiscriminately periodic) into a mixture of a periodic sinusoid in a style comparable to the Fourier series. As a result of this interactions formed simple time-invariant nonlinear simultaneous equations. Then, the solution of the nonlinear system been created by comparing the matching coefficients of the sinusoid. Let's assume that the machine started at zero voltage angle and the main and supplementary windings were $\pi/2$ electrical radians out of phase; suppose the phase voltages across the main and auxiliary windings [14] were as shown:

$$v_{as} = V_a \cos(\theta_e + \delta) \quad (26)$$

$$v_{bs} = V_b \cos(\theta_e + \delta) \quad (27)$$

where $V_a = V_b = V$, $\theta_e = \omega_e t$ and 'δ' is the voltage load angle. Therefore, the transformation used for an all-state variable in a rotor reference frame is:

$$\begin{bmatrix} F_q \\ F_d \end{bmatrix} = \begin{bmatrix} \cos(\theta_r + \varepsilon) & \sin(\theta_r + \varepsilon) \\ \sin(\theta_r + \varepsilon) & -\cos(\theta_r + \varepsilon) \end{bmatrix} \begin{bmatrix} F_a \\ F_b \end{bmatrix} \quad (28)$$

Where 'F' designates current, voltage, or flux linkage in their reference frames. $\theta_r = \omega_r t$ and 'ε' is the original rotor angular spot. Replacing the current, voltage, or flux linkage equations in their corresponding d-q rotor reference frame results in the following real complex phasor expressions with their respective coefficients as in the equations below [13, 11]:

$$F_q = \text{Re}(F_{q1} e^{j\theta_1} + F_{q2} e^{j\theta_2}) \quad (29)$$

$$F_d = \text{Re}(F_{d1} e^{j\theta_1} + F_{d2} e^{j\theta_2}) \quad (30)$$

Where, $\theta_1 = \theta_e + \theta_r = (\omega_e + \omega_r)t$
and $\theta_2 = \theta_e - \theta_r = (\omega_e - \omega_r)t$

From equations (29) and (30) the d-q voltage equations can be deduced as in equations (31) and (32).

$$V_{qs} = \text{Re}[V_{qs1}e^{j\theta_1} + V_{qs2}e^{j\theta_2}] \quad (31)$$

$$V_{ds} = \text{Re}[V_{ds1}e^{j\theta_1} + V_{ds2}e^{j\theta_2}] \quad (32)$$

where,

$$V_{qs1} = \frac{1}{2}(V_a - jV_b)e^{j\gamma_1}, \quad V_{qs2} = \frac{1}{2}(V_a + jV_b)e^{j\gamma_2} \quad (33)$$

$$V_{ds1} = jV_{qs1}, \quad V_{ds2} = -jV_{qs2} \quad (34)$$

$$\begin{bmatrix} V_{qs1} \\ V_{ds1} \\ V_{qr1} \\ V_{dr1} \\ 0 \\ 0 \\ V_{qs2} \\ V_{ds2} \\ V_{qr2} \\ V_{dr2} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} Z_{1-1} & X_{ds} & Z_{1-3} & X_{md} & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -X_{qs} & Z_{2-2} & -X_{mq} & Z_{2-4} & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ Z_{1-3} & 0 & Z_{3-3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & Z_{2-4} & 0 & Z_{4-4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2D_y & 0 & 0 & 0 & -j2\omega & -\omega & -D_y & -jD_y & 0 & 0 & 0 & 0 \\ 2D_y & 0 & 0 & 0 & \omega & -j2\omega & -D_y & jD_y & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & R_{as} & X_{ds} & 0 & X_{md} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -X_{qs} & R_{bs} & -X_{mq} & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & R_{qr} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & R_{dr} & 0 & 0 \\ -D_y & -D_y & 0 & 0 & 0 & 0 & 2D_y & 0 & 0 & 0 & 0 & -\omega \\ -D_y & jD_y & 0 & 0 & 0 & 0 & 2D_y & 0 & 0 & 0 & \omega & 0 \end{bmatrix} \begin{bmatrix} I_{qs1} \\ I_{ds1} \\ I_{qr1} \\ I_{dr1} \\ V_{qc1} \\ V_{dc1} \\ I_{qs2} \\ I_{ds2} \\ I_{qr2} \\ I_{dr2} \\ V_{qc2} \\ V_{dc2} \end{bmatrix} + \begin{bmatrix} \omega\lambda_{m1} \\ j2\omega\lambda_{m1} \\ 0 \\ j2\omega\lambda_{m1} \\ 0 \\ 0 \\ \omega\lambda_{m2} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (36)$$

In the equation (36) above,

$$D_y = \frac{1}{(4C)}, \quad Z_{1-1} = r_{aa} + j2X_{qs}, \quad Z_{2-4} = j2X_{md},$$

$$Z_{1-3} = j2X_{mq},$$

$$Z_{2-2} = r_{bb} + j2X_{ds}, \quad Z_{3-3} = r_{qr} + j2X_{qr}, \quad Z_{4-4} = r_{dr} + j2X_{dr}$$

3.1 Harmonic Balance Torque Expression

Applying phasor analysis to the electromagnetic torque equation (21) in a decoupled arrangement by replacing equation (10 - 13) in equations(1 - 9) also in decoupled form and other physical variables as expressed in equation (31) and (32) and solving algebraically, the harmonic balance torque equations expression in the parts would therefore be derived as illustrated below[14, 15].

$$T_{em} = \frac{p}{4} \text{Re} \begin{bmatrix} [(L_{md1}I_{qs1}(I_{dr1} + I_{ds1}) - L_{mq1}I_{ds1}((I_{qr1} + I_{qs1}))] \\ [(L_{md2}I_{qs2}(I_{dr2} + I_{ds2}) - L_{mq2}I_{ds2}((I_{qr2} - I_{qs2}))] \\ [(L_{md1}I_{qs2}(I_{dr1} + I_{ds1}) + L_{md2}I_{qs1}((I_{qr2} - I_{qs2})) \\ + (\lambda_{m1}I_{qs1}e^{j2\theta_1}) + (\lambda_{m2}I_{qs2}e^{j2\theta_2}) + \\ (\lambda_{m1}I_{qs2} + \lambda_{m2}I_{qs1})e^{j(\theta_1+\theta_2)}] \end{bmatrix} \quad (37)$$

where,

$$\gamma_1 = \varepsilon + \delta_0, \quad \gamma_2 = \varepsilon - \delta_0 \quad (35)$$

Following the same procedure, other state variables (current and flux linkages) are also decoupled and deduced as the current and voltage expressions. At unchanging conditions, i.e. $\omega_e = \omega_r$, and then equating to zero the derivatives of the entire new state variables, equations (1-9, 10-13, 31-34) generate the following coefficient matrix as expressed in equation (36) and the machine torque model, while the solution of equations(1-9, 10-13, and 20 - 25) using MATLAB/SIMULINK, yield the machine transient and dynamic conduct of the motor.

At synchronous speed, $\omega_r = \omega_e = \frac{d\theta}{dt}$, $\theta_1 = \theta_r + \theta_e$, also $\theta_2 = \theta_r - \theta_e$

Therefore,

$$T_{em} = \frac{p}{4} \text{Re} \left[\begin{matrix} T_1 e^{j4\theta} + T_2 + T_3 e^{j2\theta} \\ + T_4 e^{j4\theta} + T_5 + T_6 e^{j2\theta} \end{matrix} \right] \quad (38)$$

$$T_{1R} = \frac{p}{4} \text{Re} \left[\begin{matrix} (L_{md1}I_{qs1}(I_{dr1} + I_{ds1})) \\ - L_{mq1}I_{ds1}((I_{qr1} + I_{qs1})) \end{matrix} \right] e^{j2\theta_1} \quad (39)$$

$$T_2 = \frac{p}{4} \text{Re} \left[\begin{matrix} (L_{md2}I_{qs2}(I_{dr2} + I_{ds2})) \\ - L_{mq2}I_{ds2}((I_{qr2} - I_{qs2})) \end{matrix} \right] \quad (40)$$

$$T_{3S} = \frac{p}{4} \text{Re} \left[\begin{matrix} (L_{md1}I_{qs2}(I_{dr1} + I_{ds1})) \\ + L_{md2}I_{qs1}((I_{qr2} - I_{qs2})) \end{matrix} \right] e^{j(\theta_1+\theta_2)} \quad (41)$$

$$T_4 = \frac{p}{4} \text{Re} \lambda_{m1} I_{qs1} e^{j2\theta_1} \quad (42)$$

$$T_5 = \frac{p}{4} \text{Re} \lambda_{m2} I_{qs2} \quad (43)$$

$$T_6 = \frac{p}{4} \text{Re} (\lambda_{m1}I_{qs2} + \lambda_{m2}I_{qs1}) e^{j(\theta_1+\theta_2)} \quad (44)$$

$$T_{AV} = T_2 + T_5 = \frac{p}{4} \text{Re} \left[\begin{matrix} (L_{md2}I_{qs2}(I_{dr2} + I_{ds2})) \\ - L_{mq2}I_{ds2}((I_{qr2} - I_{qs2})) + \lambda_{m2}I_{qs2} \end{matrix} \right] \quad (45)$$

4. RESULTS

Fractional (0.5hp) horsepower IPM motor each for 230V, 60Hz, 8-pole, capacitor start, capacitor run IPM motor 1- Φ was employed for the research. The pertinent design data in Table 1, as well as the dynamic performance and synchronous operations of the machine, are therefore as presented in [16].

Table 1: The machine parameters for single phase 0.5 hp IPM Motor [16].

s/no	Description	Symbol	Value
1	Stator main winding resistance	ras	27.7 Ω
2	Stator auxiliary winding resistance	rbs	27.7 Ω
3	Stator d-axis leakage reactance	Llqs	67 m Ω
4	Stator q-axis leakage reactance	Llds	67 m Ω
5	Stator d-axis reactance	Lds	287 m Ω
6	Stator q-axis reactance	Lqs	511 m Ω
7	Rotor d-axis resistance	rdr	11 Ω
8	Rotor q-axis resistance	rqr	25.8 Ω
9	Rotor d-axis leakage reactance	Lldr	178 m Ω
10	Rotor q-axis leakage reactance	Llqr	179 m Ω
11	Magnetizing d-axis reactance	Lmd	220 m Ω
12	Magnetizing q-axis reactance	Lmq	441 m Ω
13	Rotor damper d-axis reactance	Ldr	378 m Ω
14	Rotor damper q-axis reactance	Lqr	622 m Ω
15	Rotor moment of inertia	J	0.0023 Kg m^2
16	Magnetic flux linkage	λ_m	0.4 Wb.turn
17	Starting Capacitor	Cs	50 μF
18	Running Capacitor	Cr	15 μF

4.1 The Motor Speed

Figure 3 showed the dynamic buildup speed from the initial value to synchronous speed for the single-phase IPM motor. The starting of the IPM motor was much more difficult due to the magnetic braking torque and backward moving negative torque. The run-up speed curves showed substantial oscillatory speed response from about 50 radians/sec to 340 radian/sec as a result of magnet excitation and instantly, it ran up to synchronism after a small overshoot of above 400 radians/sec. It also shows the responses of the motor speed after receiving the step Load of 1N at a step time of 1 second. The motor speed responses with time progressively increased in oscillation as the step load increased. The intense oscillatory torque of the IPM motor can be broadly highlighted in Figure 4, toward the mid-portion of torque-speed response where it was puffed up; over

few cycles, the motor violently oscillated, it was able to attain synchronism at 375 radian/sec as displayed.

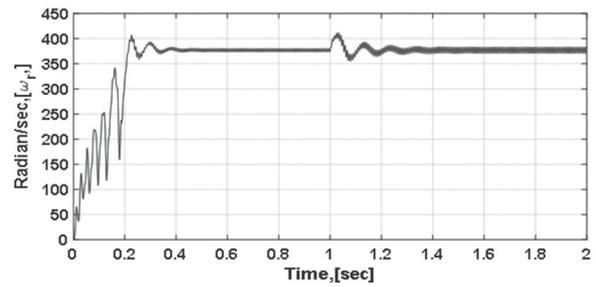


Fig. 3. Speed Versus Time of IPM Motor on 1N

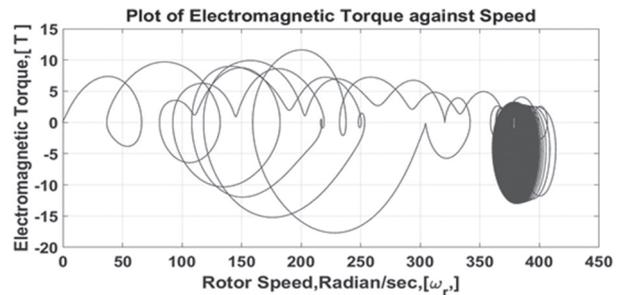


Fig. 4. Torque (T_{em}) versus Speed for IPM Motor.

4.2 Motor Currents:

Figure 5(a) and (b) showcased the core and auxiliary winding of stator currents. In the plots of the stator main and auxiliary windings, it can be easily noted that the magnet braking torque has more impact on the main windings' component than the auxiliary windings since the current waveform displayed amplitude of 2.2 A and 1.5 A at a step load of 1N respectively. In Figure 6, the readings of main and auxiliary winding currents on varying loads were displayed. Figure 7 illustrates the plots of electromagnetic, cage, and magnet torque on varying load against load angle which showed readings of 4.7 Nm, 3.25 Nm, and -7 Nm respectively. Theoretically, the summation of cage torque with the magnet torque resulted in the electromagnetic torque, that is, $T_{em} = T_c + T_m$. The bond among the torque plots of Figure 7 slightly did not conform to the theory; the minor difference in the peak torque values of the plots as can be observed accounts for the saturation effect which was not taken into account in the dynamic machine equations.

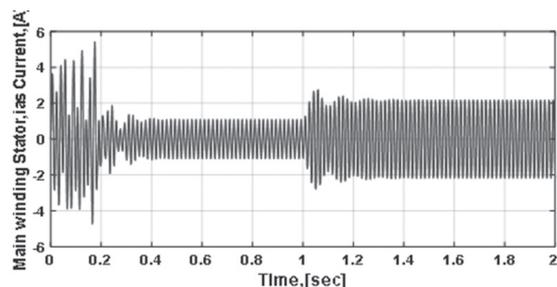


Fig. 5.(a) Current (I_{as}) Plots and Time

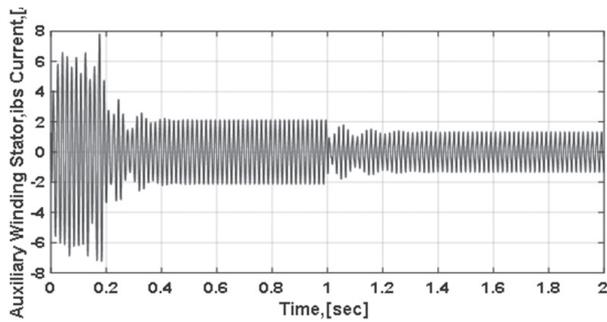


Fig. 5.(b) Current (lbs) Plots Versus Time

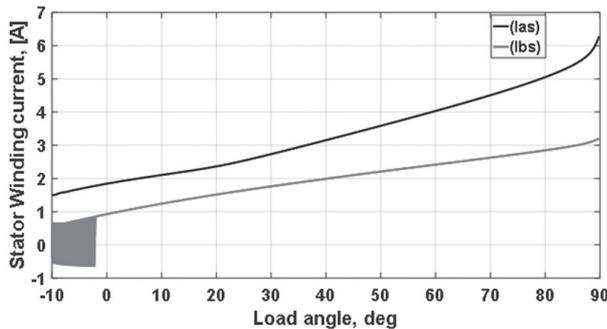


Fig. 6. Main and Aux.Winding plots against Load Angle

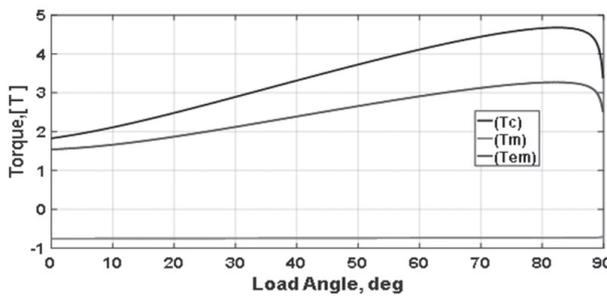


Fig. 7. Torque Plots against Load Angle

4.3 Harmonic Balance Results

This paper presented various phasor plots of stator main and auxiliary winding currents in Figure 8, q-d stator currents in Figure 9 were plotted with time, the capacitor voltage in Figure 10, and rotor currents as showcased in Figure 11; all these plotted were in response to time. The torque plots were presented in Figures 12, 14, and 16, all were plotted against time at 42° load angles while Figures 13, 15, and 17 were the responses of torques plots at a varying load angle. The phasor current plots of Figure 8 gave the closest values for single-phase characteristics which to a great extent conform to the simulated result. It can be verified in Figure 17 that the range of the load angles at which the motor can operate with minimal torque pulsation was between 40° to 50° (electrical); since in the plot of T3S, the range at the point of operating load angle was the closest range to zero pulsation.

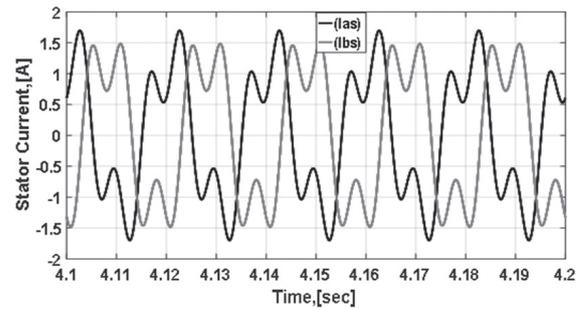


Fig. 8. Current Plots Versus Time

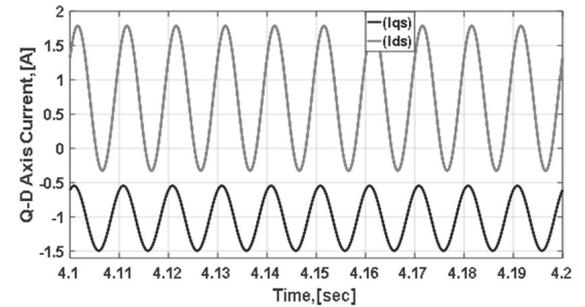


Fig. 9. Stator Current Plots against Time t

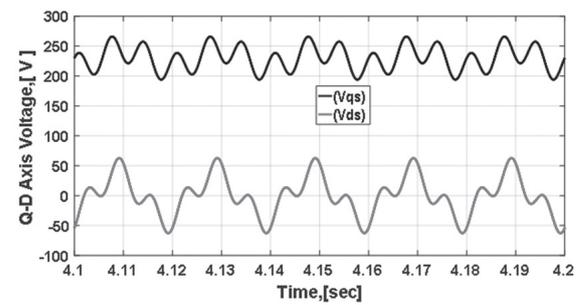


Fig. 10. Capacitor Voltage Plots Versus Time

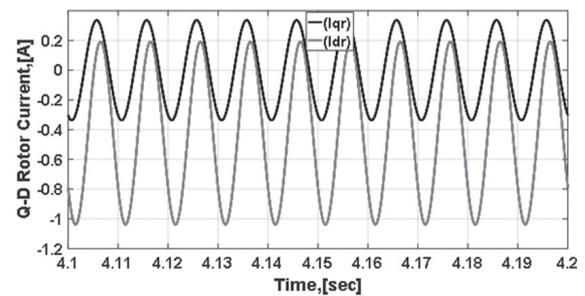


Fig. 11. Rotor Current Plots against Time

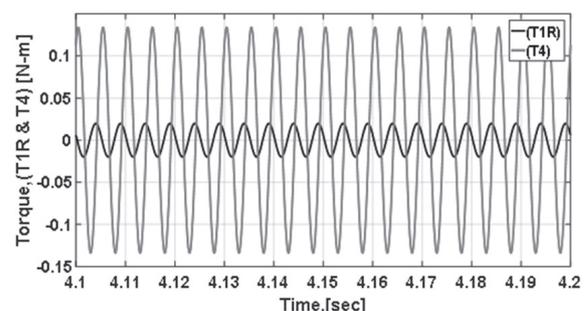


Fig. 12. Pulsating Torque Plots Versus Time

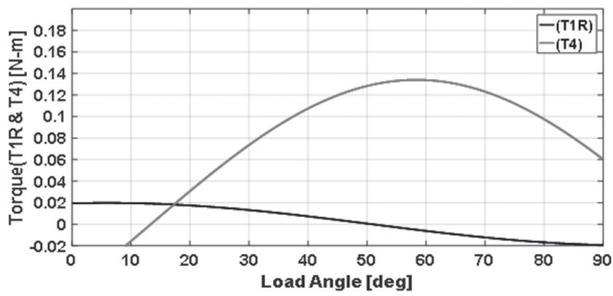


Fig. 13. Pulsating Torques Plots against Load Angle

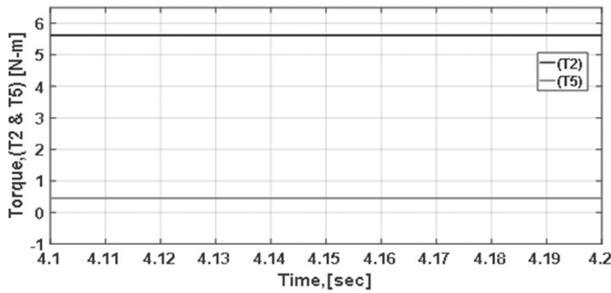


Fig. 14. Average Torque Plots Versus Time

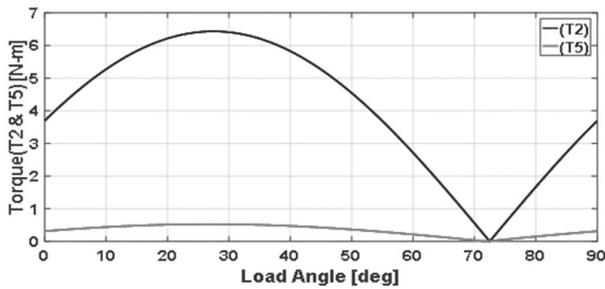


Fig. 15. Average Torques Plots against Load Angle

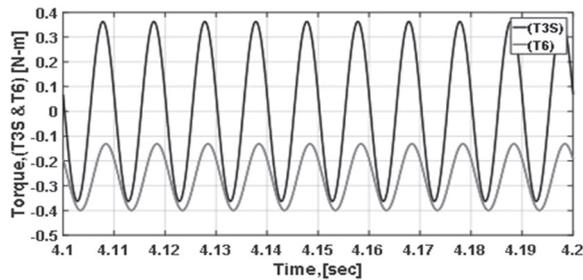


Fig. 16. Pulsating Torque Plots Versus Time

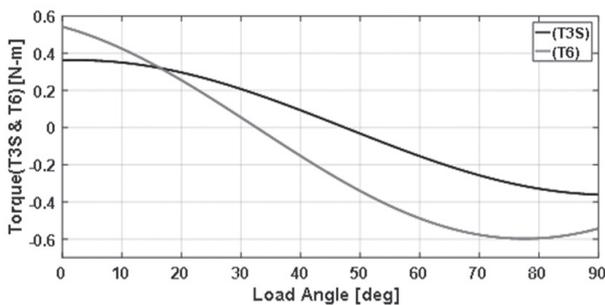


Fig. 17. Pulsating Torques Plots against Load Angle

4.4 Discussion of Results

The various capacitor voltages, currents, and torque components and their responses to time and load angles both for the dynamic simulation and the computations are as displayed above in Figures 3 - 17. The motor torques of the computed results ranging from average to pulsating torques were also displayed. There are two number average torque components while the pulsating torques are four in number. The major contributor to torque pulsation is the double frequency component (T3S & T6) as can be observed. The total indicator runout (TIR) was a pulsating torque component that has a quadruple frequency due to the rotor saliency but oscillating at $(\omega_r + \omega_e)$ and $(\omega_r - \omega_e)$; this arose as a result of the rotor saliency. It is relying on the weak capacitor value and cannot be curtailed; their minimization hinged on the rotor geometry and was gotten from the ratio of L_d/L_q of the machine, the higher the ratio the healthier the machine proficiency and power factor. T_2 was the arithmetic mean of the torque component that was hooked up on both capacitor value and the voltage load angle (δ). Their dependency on capacitor value is strong. If for any reason the rotor speed swerved from the synchronous speed, then it can dangle at $(\omega_r - \omega_e)$, meanwhile, as $\omega_e = \omega_r$ (at steady state), it produced an average torque component. Furthermore, T_{3S} was a dual-frequency aspect that oscillated at $2\omega_e$ and arose as a result of disparity in the magnetomotive force in the stator rising from main and auxiliary windings. Its reliance on capacitor was too tight hence could be curtailed. T_4 was the quadruple element established as a result of the collaboration within the onward and regressive components of the permanent magnet flux. The signal was weak and the minimization made an insignificant influence on the entire torque performance. T_5 was another average torque that boosted the machine's performance. T_6 was a twofold frequency constituent that oscillated at $2\omega_e$ and ascended as a result of an imbalance in the rotor permanent magnet flux. It was a negative revolving torque component beyond 330 load angle but from 0° to 20° , it revolved in the positive direction, unlike other pulsating torque components that were independent of load angle (δ) of the machine and oscillate in the negative and positive direction at any point in time.

5. CONCLUSIONS

Single-phase synchronous motor analysis is not easy to come by owing to its inherent asymmetrical features, but it even becomes much complicated with the introduction of a permanent magnet as an integral part of it, since permanent magnet machine performance indices are characterized by high torque ripples. However, the painstaking efforts put in this novelty work for the IPM machine analysis have successfully paid off. This paper showcased a unified study of high-efficiency single-phase line-start IPM motor that permitted the obtaining of the steady-state, as well as dynamic and

transient performances analysis of the motor for different functional settings. A mathematical model developed for the motor was based on d-q axis theory and harmonic balance techniques, coupled with the derivation of the inductances as peculiar to the machine. Simulation and computed results of the fractional horsepower having optimum starting and running stator capacitance of $50\mu F$ and $15\mu F$ for different operating conditions are presented.

The simulated results to a large extent conform with computed results; the minor variation could be ascribed to the non-enclosure of saturation effect of the third harmonic frequency components inherent in the stator of single-phase machines in the model equation of the dynamic simulation. Most importantly, the work has discovered the average and pulsating torques inherent in IPM machines during its asynchronous and steady-state mode of operations. Those torques as discovered and presented have been categorized and their functional pulsating characteristics fully analyzed. Secondly, it optimized the performance of single-phase IPM machines as the impact of the pulsating torques can be extensively minimized by operating the machine within the range of the load angles at which the motor can operate with minimal torque pulsations [17]; which is between 40° to 50° (electrical) since in T3S, that range at the point of operating load angle of 42° is the closest range to zero pulsation. At this point, it becomes imperative that a narrow load angle range as this can not be the best expected. Hence, further effort should be made to widen the load angle gap, incorporate electronic circuits for better control, taking their attendant setbacks into consideration.

6. REFERENCES

- [1]. M. Popescu, T. J. E. Miller, M. I. Mcgilp, G. Strappazon, N. Trivillin, R. Santarossa, "Line-start permanent-magnet motor: Single-phase starting performance analysis", IEEE Transactions on Industry Applications, Vol. 39, No. 4, 2003, pp. 1021-1030.
- [2]. T. J. E. Miller, "Single-phase permanent-magnet motor analysis", IEEE Transactions on Industry Applications Vol. IA-2 1, No. 3, 1985, pp. 651-658.
- [3]. O. Ojo, A. Ginart, O. Omozusi, A. A. Jimoh, "Modeling and analysis of a single-phase synchronous-reluctance machine including saturation effect", Conference Record of the 1997 IEEE Industry Applications Conference Thirty-Second IAS Annual Meeting, New Orleans, LA, USA, 5-9 October 1997, pp. 290-301.
- [4]. O. A. Osaloni, "Performance analysis and nonlinear feedback control of interior permanent magnet synchronous motor", Tennessee Technological University Faculty of the Graduate School, PhD Thesis, 2003.
- [5]. B. A. Welchko, T. M. Jahns, S. Hiti, "IPM Synchronous Machine Drive Response to a Single-phase Open Circuit Fault", Proceedings of the Applied Power Electronics Conference and Exposition, Anaheim, CA, USA, 4-8 March 2001, pp. 421-427.
- [6]. O. Ojo, Z. Wu, "Modeling of the doubly-fed reluctance machines using harmonic balance technique and Manley-Rowe power frequency relationships", Conference Record of the 1996 31st IAS annual meeting, Vol2, San Diego, CA, USA, 6-10 October 1996, pp. 794-800.
- [7]. E. S. Obe, A. Binder, "Calculation of Rotor Currents of Single-phase Synchronous Reluctance Motor Using d-q Harmonic Balance Techniques", Electric Power Components and Systems, Vol. 37, No. 9, 2009, pp. 943-956.
- [8]. T. J. E. Miller, "Single-phase permanent-magnet motor analysis", IEEE Transactions on Industry Applications, Vol. IA-2 1, No. 3, 1985, pp. 651-658.
- [9]. J. W. Finch, P. J. Lawrenson, "Asynchronous performance of single-phase reluctance motors", Proceedings of the Institution of Electrical Engineers, Vol. 126, 1979, pp. 1249-1254.
- [10]. T. A. Lipo, "Performance Calculations of a Reluctance Motor Drive by DQ Harmonic Balance", IEEE Transactions on Industry Applications, Vol. IA-15, No. 1, 1979, pp. 25-35.
- [11]. M. A. Rahman, A. M. Osheiba, K. Kazumi, M. A. Jebbar, H. W. Ping, K. Wang, M. Z. Hossan, "Advances on Single-Phase Line-Start High-Efficiency Interior Permanent Magnet Motors", IEEE Transaction on Industrial Electronics, Vol. 59, No. 3, 2012, pp. 1333-1345.
- [12]. J. W. Finch, P. J. Lawrenson, "Synchronous performance of a single-phase reluctance machines", Proceedings of the Institution of Electrical Engineers, Vol. 125, No. 12, 1978, pp. 1350-1356.
- [13]. R. Kerkman, "Machine analysis with unbalanced terminal constraints by d-q harmonic balance", IEE Proceedings B (Electric Power Applications), Vol. 128, 1981, pp. 343-357.

- [14]. E. S. Obe, O. Ojo "Line-start performance of single-phase synchronous reluctance motor with a controlled capacitor", IEE Proceedings - Electric Power Applications, Vol. 152, No. 4, 2005, pp. 967-976.
- [15]. E. S. Obe, T. Senjyu, M. U. Agu, L. U. Anih, O. Ojo, "Steady-state components in a balanced single-phase synchronous reluctance motor with a running capacitor", Electric Power Systems Research, Vol. 77, No. 3-4, 2007, pp. 295-302.
- [16]. M. A. Rahman, A. M. Osheiba, K. Kazumi, M. A. Jebbar, H. W. Ping, K. Wang, M. Z. Hossan, "Advances on Single-Phase Line-Start High-Efficiency Interior Permanent Magnet Motors", IEEE Transaction on Industrial Electronics, Vol. 59, No. 3, 2012, pp. 1333-1345.
- [17]. M. J. Thomas, W. L. Soong, "Pulsating Torque Minimization Techniques for Permanent Magnet AC Motor Drives-A Review", IEEE transactions on industrial electronics, Vol. 43, No. 2, 1996, pp. 321-329.

Battery energy storage technologies overview

Review paper

Zvonimir Šimić

J. J. Strossmayer University of Osijek,
Faculty of Electrical Engineering,
Computer Science and Information Technology Osijek
Kneza Trpimira 2b, Osijek, Croatia
zvonimir.simic@ferit.hr

Danijel Topić

J. J. Strossmayer University of Osijek,
Faculty of Electrical Engineering,
Computer Science and Information Technology Osijek
Kneza Trpimira 2b, Osijek, Croatia
danijel.topic@ferit.hr

Goran Knežević

J. J. Strossmayer University of Osijek,
Faculty of Electrical Engineering,
Computer Science and Information Technology Osijek
Kneza Trpimira 2b, Osijek, Croatia
goran.knezevic@ferit.hr

Denis Pelin

J. J. Strossmayer University of Osijek,
Faculty of Electrical Engineering,
Computer Science and Information Technology Osijek
Kneza Trpimira 2b, Osijek, Croatia
denis.pelin@ferit.hr

Abstract – Battery technologies overview for energy storage applications in power systems is given. Lead-acid, lithium-ion, nickel-cadmium, nickel-metal hydride, sodium-sulfur and vanadium-redox flow batteries are overviewed. Description, graphical representation, advantages and disadvantages as well as technical characteristics are given for all technologies. Differences and similarities between different battery technologies are perceived. Battery technologies are considered with respect to peak shaving, load leveling, power reserve, integration of renewable energy, voltage and frequency regulation and uninterruptible power supply applications. According to technical characteristics for overviewed technologies, comparison between battery storage technologies is given through diagrams which are uniformed. Comparison is done according to specific power, specific energy, power density, energy density, power cost, energy cost, lifetime, lifetime cycles, cell voltage and battery technology efficiency.

Keywords – battery, technologies, overview, energy storage, applications, power system, characteristics

1. INTRODUCTION

There is a strong growth in the development of energy storage technologies, especially batteries, in the last decades. This is confirmed by the large number of available papers in which the overview of storage technologies and their applications is given. Some storage technologies are in use for a long time while other technologies are under research and development. Some technologies are widely used in commercial applications while others are used only for experimental applications.

Energy storage is important for matching electricity supply to load demand, increasing power quality and enabling renewable technologies integration [1]. Selection of suitable energy storage technology depends on power and energy capacity, the period during which energy needs to be available on the grid, costs, space required for technology placement and location in the network [2]. Power systems in the future could not be possible without storage. Pumped hydro storage was dominant throughout history, batteries are suitable for load shifting while flywheel, capacitors and superconductive inductors are suitable for voltage regulation

[3]. Batteries are an ideal technology for continuous energy storage applications, while flywheels and super capacitors are ideal for power storage applications [2]. For large scale energy storage applications, the most suitable technology is a pumped hydro because of technical maturity [4]. Batteries are the cheapest technology for different applications, flywheel is suitable for short storage periods while compressed air and pumped hydro are suitable for large scale applications [5]. In comparison to batteries and supercapacitors, using hybrid storage technologies lead to reduction of volume, weight and costs of storage systems [1].

Battery energy storage system (BESS) functionalities can be classified at different grid levels as following: generation, transmission and distribution, end-user and renewable energy sources (RES) integration [6]. Batteries are ideal for energy storage applications because of the short response time, modularity, flexible installation and short construction time. Batteries can meet the requirements of Grid-level large-scale electrical energy storage (GLEES) [7]. There are four main groups in which batteries can be classified: primary cells, secondary cells, reserve batteries and fuel cells. Secondary cells imply rechargeable batteries such as

lead-acid batteries, alkaline batteries, lithium batteries, sodium batteries, flow batteries and other batteries [8]. There is a problem when batteries work in a hard environment, especially in low temperatures which can be overcome with implementation of specific materials, but a more acceptable solution is using the combination of existing battery technologies [9]. The weather impact on renewables, specifically photovoltaics, can be overcome with the use of energy storage systems. Battery storage technologies, unlike other storage technologies, are more suitable for renewable energy sources because of a simple and efficient way of electrical energy storing [10]. Lithium-ion and lead-acid are most suitable for RES integration applications [6]. Small scale batteries at consumption level have ability to solve photovoltaic variability problems. On the other side, for wind generation, additional optimization of a storage system is required. Designing a storage system is a complex process, but it is possible to make effective solutions for many different applications [11].

Lead-acid batteries are in use since a long time ago and all technology limits and possibilities are well known and like other battery technologies, lead-acid batteries are suitable for different grid applications [12]. Lithium-ion technology is widely used in small scale portable applications because of high power and energy and low weight but it is expensive for large scale stationary applications [4]. Lithium-ion battery technology is mostly used in electric vehicles because of a high power and energy. Main obstacles for using lithium-ion batteries for stationary applications are battery aging and efficiency drop due to aging. There is a possibility of second use batteries for stationary applications [13]. Lithium-ion battery technology has a large potential to become one of the main technologies for grid applications because of a high efficiency, energy density and long cycle life, but to achieve this, it is required to decrease technology costs, build effective systems for collecting and recycling and develop new lithium technologies [14].

Research is focused on electrochemical storage technologies. Development of battery technologies with high efficiency and low cost is required for large scale grid applications. There are potassium-ion batteries with high energy density, but low cost in comparison to lithium-ion batteries [15]. Research and development of battery technologies in the future will be focused on solid state batteries, which are safe and have high energy density and currently used batteries with liquid electrolyte will be replaced with solid state batteries [16]. Future development of battery storage technologies is focused on improving storage system efficiency and decreasing investment cost [17]. Development of storage systems should be focused on technologies with high performance and low prices. Environmental impact of storage technologies, especially batteries, is significant [18]. There are needs for development of microgrids with energy storage systems, but using only

one storage technology cannot meet all requirements of a microgrid so there is a need for development of hybrid storage technologies for microgrid applications [19]. The number of plug-in electric vehicles is constantly growing and, in the future, electric vehicles will participate as a virtual energy storage system and this can reduce the need for stationary storage [20].

This paper deals with battery storage technologies overview and their grid applications. The paper is divided into 5 chapters. Chapter 1 gives a short overview of available papers which deals with battery storage technologies. In chapter 2, six battery technologies are explained in the following order: lead-acid, lithium-ion, nickel-cadmium, nickel-metal hydride, sodium-sulfur and vanadium-redox flow batteries. In chapter 3, following grid applications are explained: peak shaving, load leveling, power reserve, integration of renewable energy, voltage and frequency regulation and uninterruptible power supply, Grid applications are presented textually and graphically. In chapter 4, comparison of battery technologies is done through diagrams according to characteristics given in the chapter 2. Chapter 5 will give conclusions about overviewed battery technologies.

This paper will give contributions through an overview of the most used battery technologies for different grid applications. All technologies are overviewed in the same way, first textual description, then graphical representation, then advantages and disadvantages and the technology characteristics. One of the observed characteristics is energy and power cost which will be expressed in EUR, unlike most of the reviewed papers in which it is expressed in USD. Overviewed technologies will be compared through uniformed diagrams. In this paper, specific power and specific energy, then power density and energy density, then power and energy cost, then lifetime and the number of lifetime cycles, then cell voltage and the efficiency for battery technologies will be compared.

2. BATTERY TECHNOLOGIES OVERVIEW

With the development of smart grids and microgrids there is a growing need for energy storage in power systems. Throughout history various storage systems have been developed for electrical energy storage. The main difference between various storage types is in a form of energy in which electrical energy is stored. There are mechanical, thermal, thermochemical, chemical, electrochemical, electrical and magnetic storage types. Main mechanical storage technologies are flywheels, pumped hydro and compressed air. Electrical storage technologies are capacitors and supercapacitors. Main chemical storage technology are hydrogen fuel cells.

This paper deals with batteries which are an electrochemical storage technology. There are different battery technologies, and the several types will be explained. There are also flow battery technologies and

one of them will be explained. In this paper the following battery technologies are presented:

- lead-acid,
- lithium-ion,
- nickel-cadmium,
- nickel-metal hydride,
- sodium-sulfur,
- vanadium-redox flow.

The overview of listed battery technologies is done according to different technical characteristics. Observed technical characteristics are defined below.

Cell voltage [V] is a voltage measured on the battery cell between positive and negative terminals. It is very important for stacking cells into batteries with previously defined voltage.

Specific energy [Wh/kg] represents the available energy of battery technology per unit of mass. It is important for comparing the output energy of different technologies with equal mass.

Specific power [W/kg] represents the available power of battery technology per unit of mass. It is important for comparing the output power of different technologies with equal mass.

Energy density [kWh/m³] defines the available energy of battery technology per unit of volume. It is important for comparing the output energy of different technologies with equal volume.

Power density [kW/m³] defines the available power of battery technology per unit of volume. It is important for comparing the output power of different technologies with equal volume.

Efficiency [%] shows the ratio between energy that can be discharged from battery and energy used to charge battery. There are losses during energy conversion from electrical to electrochemical and then during conversion from electrochemical to electrical.

Working temperature [°C] is the temperature range at which the battery technology can operate. If temperature is out of this range it can come to significant reduction of battery performance.

Lifetime cycles [cycles] represents the number of charge and discharge processes during which the battery keeps minimum working performance. It's important for applications with a lot of charging and discharging processes because there are a high number of cycles.

Lifetime [years] represents the number of years during which the battery keeps the minimum working performance. If lifetime is short, batteries need to be replaced often, which is expensive.

Maximum depth of discharge [%] represents the amount of charge that can be discharged in one cycle. Some technologies cannot be fully discharged.

Self-discharging rate [%] represents the amount of charge that is discharged when the battery is not in use. It is not significant if battery storage is charging and discharging all the time.

Power rating [MW] shows power capacity of installed batteries on the same location which can be considered as a single storage unit. It is important for planning large scale battery storage systems.

Energy cost [€/kWh] defines the price of battery energy storage per unit of energy. Prices in EUR are converted from prices in USD.

Power cost [€/kW] defines the price of battery energy storage per unit of power. Prices in EUR are converted from prices in USD. Given prices are approximate and intended only for technology comparison.

2.1 Lead-acid

Lead-acid (Pb-acid) technology has been in use for a long time, it is easy and cheap for installation and maintenance, so this is the main reason for wide use of this technology and this technology is also one of the most common for stationary applications worldwide. Lead-acid batteries have ability to perform a deep discharge when it is required and the main problem with lead-acid batteries is that battery performance largely depends on temperature [21]. Nominal voltage of this technology cell is around 2 volts. Lead-acid battery technology is based on positive and negative electrodes submerged into electrolyte which is a combination of sulfuric acid and water, lead dioxide is used as a positive electrode and lead is used as a negative electrode [12]. Lead-acid battery technology cell is shown in Fig. 1.

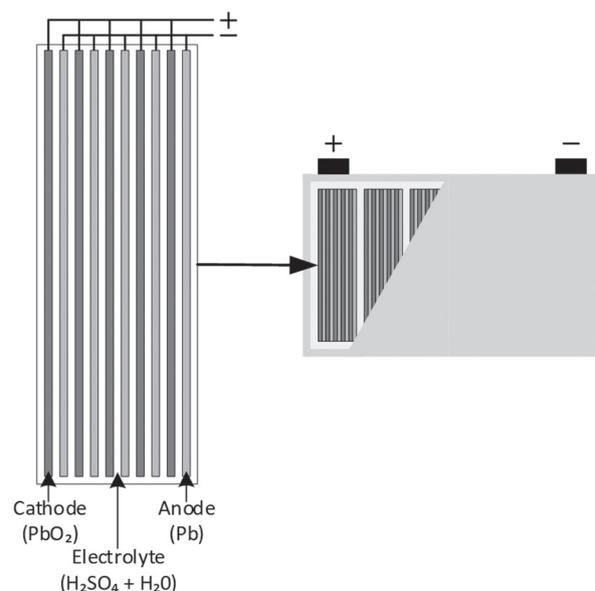


Fig. 1. Lead-acid battery cell

Some advantages of the lead-acid technology are low cost, high cell voltage, suitability for intermittent charge applications and good ability of recycling. Disadvantages are limited energy density and the number

of lifetime cycles which is lower in comparison to other technologies [7]. Detailed characteristics of the lead-acid battery technology are shown in table 1.

Table 1. Lead-acid characteristics [4, 5, 7, 8, 19]

Characteristics	Value
Cell voltage	2 – 2.1 V
Specific energy	25 – 50 Wh/kg
Specific power	150 – 400 W/kg
Energy density	25 – 90 kWh/m ³
Power density	10 – 400 kW/m ³
Efficiency	63 – 90 %
Working temperature	18 – 45 °C
Lifetime cycles	250 – 2000
Lifetime	2 - 15 years
Max. depth of discharge	80 %
Self-discharge rate	0.1 – 0.3 % per day
Power rating	0 – 20 MW
Energy cost	40 – 170 €/kWh
Power cost	250 – 500 €/kW

2.2. Lithium-ion

Lithium-ion (Li-ion) technology is one of the most advanced battery technologies widely used today. Cell-phones, smartphones, tablets, laptops, all gadgets are powered with the Li-ion battery. There are many pros of the lithium-ion technology: high power, energy capacity, long battery lifetime and relatively low weight and because of these pros, Li-ion technology is being used to power hybrid and electric vehicles [22]. Lithium-ion battery working principle is given in Fig. 2 according to [13]. Li-ion cells consist of two electrodes, anode and cathode. Graphite is used as anode and the lithium metal oxide is used as cathode. The lithium salt in organic solvent is used as electrolyte. Anode collector is made from copper (Cu) and cathode collector is made from aluminum (Al). Working principle of this technology is based on Li-ions moving from cathode to anode when battery is in charging process and from anode to cathode when battery is in discharging process [23].

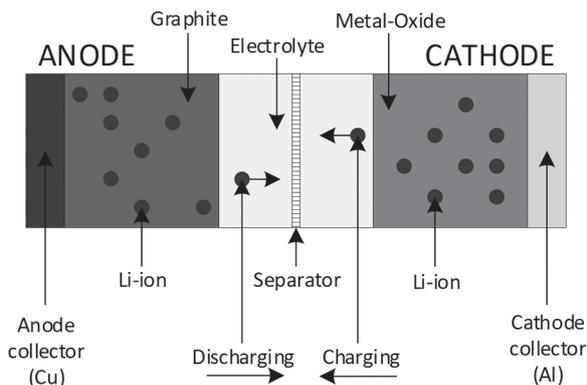


Fig. 2. Lithium-ion technology working principle

Advantages of the lithium-ion technology are long battery lifetime, the number of lifetime cycles, high energy density, low maintenance cost and there is no memory effect. Disadvantages are high cost in comparison to other technologies, poor performance at high temperature and the request for protective circuits [7]. Detailed characteristics of the lithium-ion battery technology are shown in table 2.

Table 2. Lithium-ion characteristics [4, 5, 7, 8, 19]

Characteristics	Value
Cell voltage	2.5 – 5 V
Specific energy	80 – 250 Wh/kg
Specific power	200 – 2000 W/kg
Energy density	95 – 500 kWh/m ³
Power density	50 – 800 kW/m ³
Efficiency	75 – 97 %
Working temperature	20 – 65 °C
Lifetime cycles	100 – 10000
Lifetime	5 - 15 years
Max. depth of discharge	100 %
Self-discharge rate	0.1 – 0.3 % per day
Power rating	0 – 0.1 MW
Energy cost	500 – 2100 €/kWh
Power cost	1000 – 3400 €/kW

2.3. Nickel-cadmium

Nickel-cadmium (Ni-Cd) technology is in use for a long time in applications that require a long battery life and in difficult environmental conditions because this battery technology is cheap and robust. Nickel-cadmium technology is based on cathode made from nickel oxide hydroxide and anode made from metallic cadmium while electrolyte used for Ni-Cd batteries is potassium hydroxide [24]. Cross section of the nickel-cadmium battery is shown in Fig. 3 according to [4]. Ni-Cd batteries can be charged with a high charge rate which means that the battery is charging with a current much higher than a nominal current is, but in this case the charging process must be stopped when the battery is full, otherwise the battery will heat very fast which leads to damage. The main problem with Ni-Cd batteries is a memory effect which means that battery loses full capacity if it is slightly discharging and recharging every time during a certain period [25].

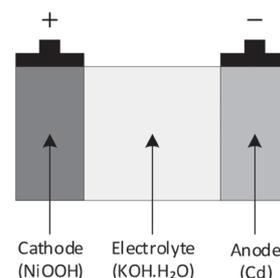


Fig. 3. Nickel-cadmium battery cross section

Advantages of NiCd technology are low maintenance cost, the number of lifetime cycles, suitability for long-term storage and ability to resist electrical and physical stress. Disadvantages are high cost in comparison to the lead-acid technology, limited energy density, toxic and caustic elements in batteries and the memory effect [7]. Detailed characteristics of NiCd battery technology are shown in table 3.

Table 3. NiCd characteristics [4, 5, 7, 8, 19]

Characteristics	Value
Cell voltage	1.2 – 1.3 V
Specific energy	30 – 80 Wh/kg
Specific power	80 – 300 W/kg
Energy density	15 – 150 kWh/m ³
Power density	40 – 140 kW/m ³
Efficiency	60 – 90 %
Working temperature	- 40 – 50 °C
Lifetime cycles	1000 – 5000
Lifetime	10 - 20 years
Max. depth of discharge	80 %
Self-discharge rate	0.2 – 0.6 % per day
Power rating	0 – 40 MW
Energy cost	680 – 1300 €/kWh
Power cost	420 – 1300 €/kW

2.4. Nickel-metal hydride

Nickel-metal hydride (Ni-MH) technology has been used in several applications such as energy storage for smart energy systems, robust battery systems which work at high temperatures, hybrid electric cars and public transport [26]. Ni-MH battery cell cross section with the main parts is shown in Fig. 4 according to [27]. Nickel-metal hydride technology is based on the negative electrode made from hydrogen-absorbing alloys which have the possibility to absorb releasing hydrogen and the positive electrode made from nickel oxy-hydroxide. There is a separator which separates positive and negative electrodes to prevent shorting between electrodes. Electrolyte used in this technology is potassium hydroxide (KOH). There is a current collector made of metal which minimizes the internal battery resistance. To release gases produced during the overcharging or shorting there is a self-sealing safety vent [28].

Some advantages of NiMH technology are long battery lifetime, high number of lifetime cycles, good performance at the high temperatures, high energy density, good ability of recycling and the high tolerance to battery overcharging and over discharging. Disadvantages are high cost in comparison to the lead-acid technology and bad performance at the low working temperatures [7]. Detailed characteristics of NiMH battery technology are shown in table 4.

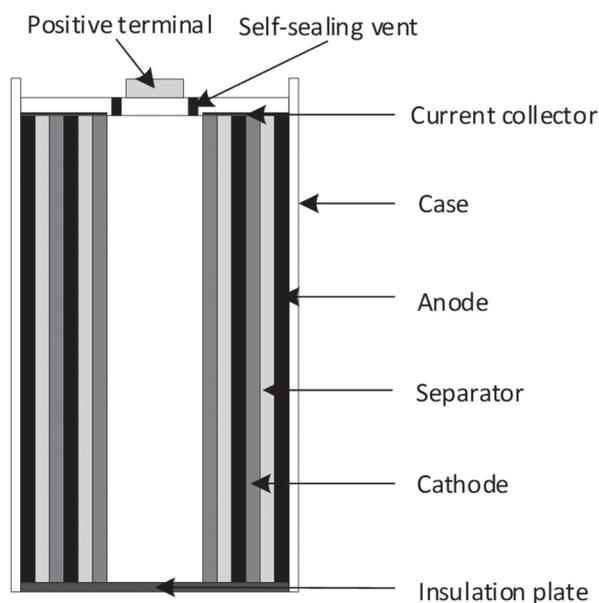


Fig. 4. Nickel-metal hydride battery cross section

Table 4. NiMH characteristics [4, 5, 7, 8, 19]

Characteristics	Value
Cell voltage	1.2 – 1.35 V
Specific energy	40 – 110 Wh/kg
Specific power	200 – 300 W/kg
Energy density	40 – 300 kWh/m ³
Power density	10 – 600 kW/m ³
Efficiency	50 – 80 %
Working temperature	- 30 – 70 °C
Lifetime cycles	300 – 1800
Lifetime	2 - 15 years
Max. depth of discharge	100 %
Self-discharge rate	5 – 20 % per day
Power rating	0.01 – 3 MW
Energy cost	170 – 640 €/kWh
Power cost	200 – 470 €/kW

2.5. Sodium-sulfur

Sodium-sulfur (NaS) battery technology is one of the most suitable for use in energy storage systems because of the high energy density [29]. NaS battery cell cross section is shown in Fig. 5 according to [29]. This technology is based on the use of sodium as anode and sulfur as cathode, electrolyte is beta alumina ceramics. For this technology is interesting that electrolyte is also a separator. NaS technology has a low internal cell resistance because of the use a ceramic electrolyte and this is good for two reasons. With the low resistance, power to weight ratio is being increased and heat produced during the charging process is being decreased [30]. Typically working temperature of the sodium-sulfur

batteries is between 300°C and 350°C. Reason for high temperature is to keep electrodes in a liquid state. High temperature decreases efficiency of the cycle which increases the number of operation cycles. [31]

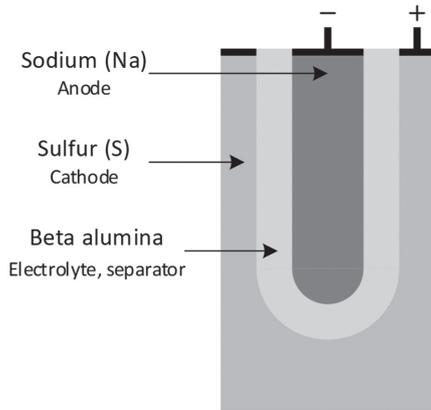


Fig. 5. Sodium-sulfur battery cell cross section

Advantages of sodium-sulfur technology are high energy density, high number of lifetime cycles, capability of pulse power and good resistivity to self-discharging. Disadvantages are high cost and high temperature required for battery operation [7]. Detailed characteristics of sodium-sulfur battery technology are shown in table 5.

Table 5. NaS characteristics [4, 5, 7, 8, 19]

Characteristics	Value
Cell voltage	1.8 – 2.71 V
Specific energy	150 – 240 Wh/kg
Specific power	90 – 230 W/kg
Energy density	150 – 350 kWh/m ³
Power density	1.2 – 50 kW/m ³
Efficiency	75 – 90 %
Working temperature	300 – 350 °C
Lifetime cycles	2500 – 40000
Lifetime	10 - 15 years
Max. depth of discharge	100 %
Self-discharge rate	0 % per day
Power rating	0.05 – 34 MW
Energy cost	250 – 420 €/kWh
Power cost	850 – 2500 €/kW

2.6. Vanadium-redox flow battery

Vanadium-redox flow battery (VRFB) is a new technology which promises a lot because of the very good characteristics. This technology has a long lifetime, very fast response time and long storage time which is ideal for long-term energy storage. Power and energy of VRFB are independent, power depends on the number and size of the cells and energy depends on the available electrolyte, respectively tank size [32]. Work-

ing principle of the VRFB technology is shown in Fig. 6 according to [33]. VRFB technology is based on two tanks in which vanadium ions electrolytes are stored, one electrolyte is positive and the other is negative. Flow of electrons is caused by oxidation and reduction processes in the ion selective membrane through which electrolytes are being pumped. Response time of this technology is fast because electrolyte flow does not change, regardless of whether battery is charging or discharging [33].

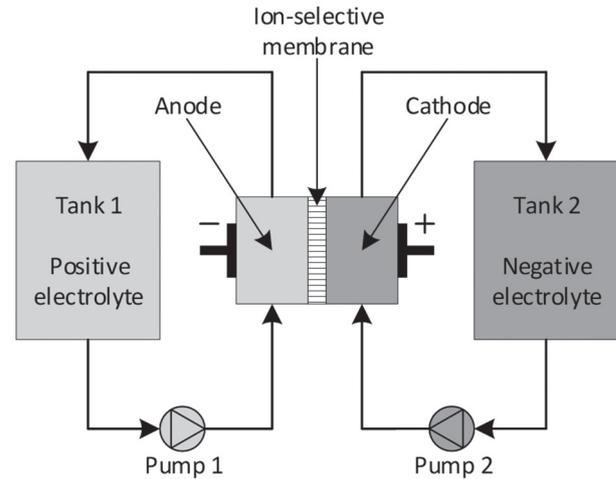


Fig. 6. Vanadium-redox flow battery working principle

Advantages of the vanadium-redox flow battery technology are high operating safety, high number of lifetime cycles, low operation and maintenance cost and deep discharging capability. Disadvantages are low energy density and the large space required for technology placement [7]. Detailed characteristics of the vanadium-redox flow battery technology are shown in table 6.

Table 6. VRFB characteristics [4, 5, 7, 8, 19]

Characteristics	Value
Cell voltage	1.2 – 1.4 V
Specific energy	10 – 130 Wh/kg
Specific power	50 – 150 W/kg
Energy density	10 – 33 kWh/m ³
Power density	2.5 – 33 kW/m ³
Efficiency	75 – 90 %
Working temperature	5 – 45 °C
Lifetime cycles	10000 – 16000
Lifetime	5 - 15 years
Max. depth of discharge	100 %
Self-discharge rate	0 % per day
Power rating	0.03 – 3 MW
Energy cost	130 – 850 €/kWh
Power cost	500 – 1300 €/kW

3. GRID APPLICATIONS OF A BATTERY STORAGE

With growth of the amount of distributed generation, passive distribution grids become active. It means that energy flows are in two ways, from generation, over transmission grid to distribution grid and in the opposite way, from distribution grid to transmission grid. Operation of an active distribution grid is more complicated than operation of a passive grid, especially if distributed generation is intermittent, such as photovoltaic and wind. There is a demand for additional energy storage which has ability to support grid functionality and stability. In conventional power systems, load profile is divided on a base load, which is covered from baseload power plants and on a variable load, which is covered from load following power plants. Variable load can be partially covered from energy storage so there is less need for regulation power plants. Batteries are more acceptable for voltage and frequency regulation application than regulation power plants because of the fast response on voltage and frequency changes. Batteries are being used in the following grid applications:

- peak shaving,
- load leveling,
- power reserve,
- integration of renewable energy sources,
- voltage and frequency regulation,
- uninterruptible power supply.

3.1. Peak shaving

Highest demand for electrical energy appears in the evening during winter because of the high amount of energy is required for heating and in the middle of the day during summer because of the high amount of energy is required for cooling. Grid utility operators have introduced different prices of electricity depending on energy demand. When demand is high, price is higher and when demand is low, price is lower. Methods for reducing consumption of electrical energy when demand is at peak are called the peak shaving and it can be done in different ways. First method is switching off appliances that are not in use and installing heating systems with thermostat to reduce power demand, second method requires the installation of additional generators to achieve peak power demand and third method is using batteries to cover consumption during peak power demand [16].

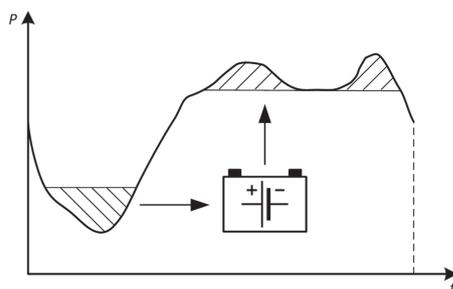


Fig. 7. Peak shaving with using of a battery storage

Peak shaving with the use of a battery storage is shown in Fig. 7. When the power demand is low, energy price is also low, and energy can be stored in batteries and when the power demand is high, energy price is also high and peak load can be partially or fully covered with energy from batteries [18].

3.2. Load leveling

Load leveling is not much different from peak shaving. Both applications have similar working principles. Energy is being stored in batteries during low power demand and being used from batteries during high power demand. Peak shaving has a task to level the load profile only during peak power demand while load leveling has a task to level the entire load profile during the entire day [18]. Load leveling uses low price energy from base generation to cover high price peak energy demand. Like the peak shaving, load leveling can give profit to the customers, because the use of stored energy during high price periods and make customers less dependent on the electricity market during peak periods [34]. Load leveling with the use of a battery storage is shown in Fig. 8.

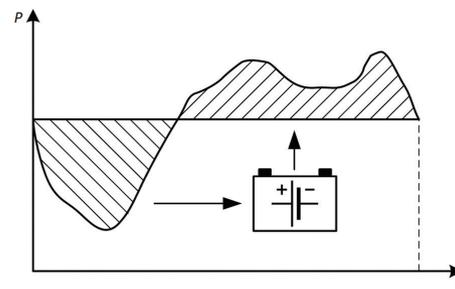


Fig. 8. Load leveling with using of a battery storage

It can be assumed that load leveling requires much higher battery capacity than peak shaving applications because peak shaving is covering consumption during only a few hours a day while load leveling is covering consumption more hours a day.

3.3. POWER RESERVE

When grid operators predict the load demand for the day ahead it can happen that actual load demand is less than predicted load demand and it also can happen that actual load demand is higher than predicted. There is a possibility to install a battery storage system that will store energy excess when the actual load demand is less than predicted and cover actual load demand when it is higher than predicted load demand as shown in Fig. 9. Battery storage systems are ideal for this application because of the fast response time [16].

Reserve in the power system is important because reserve supports the power system from the unaware load reducing when load demand is high. Battery storage is a much better solution for spinning reserve than synchronous generators which must be synchronized with the grid during startup and this makes them slow.

More reserve capacity leads to better system reliability, but it is necessary to find the optimal reserve capacity due to high installation price. In the future, electric vehicles connected to the charging stations will also be able to provide reserve in the power system [18]. Providing a reserve service to the grid operator has a financial benefit for battery storage owners. Grid operators and reserve providers enter a contract that defines the amount of energy which must be available to feed into the power system and the energy price at which the operator will buy energy from the reserve provider.

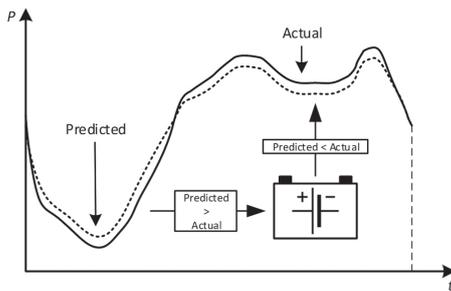


Fig. 9. Battery power reserve

3.4. Integration of renewable energy sources

Renewable energy sources nowadays are present worldwide and their share in total production of electrical energy is in constant growth. Integration of renewable energy sources, especially wind and solar, causes economical and technical challenges for grid operators [5]. Renewable energy sources are variable, intermittent and unpredictable so they are harder to integrate in the power system [35]. It is more difficult to operate the power system and keep stability because of power fluctuations [36]. Intermittency and variability of the renewable energy sources are primarily caused by weather changes. In the case of photovoltaics, it often happens that clouds reduce solar radiation which leads to decreased generation from photovoltaics. In the case of wind power, wind speed is variable which leads to constant variability in generation from wind.

Intermittency and variability can be compensated with implementation of a battery storage system which is necessary for the off-grid photovoltaic and wind power. On the other side, for on-grid photovoltaic and wind power, a battery storage is welcome because it allows integration of higher amounts of distributed generation. During power outages caused by weather changes consumption can be covered from energy stored in batteries which are distributed in the power system. As the amount of distributed generation will be increased, capacity of installed battery storage also needs to be increased because for higher amounts of solar and wind generation, more battery storage will be needed to cover consumption during power outages. Fig. 10. shows the case in which energy excess produced from photovoltaic during sunny days can be used later in the evening to cover higher energy demand.

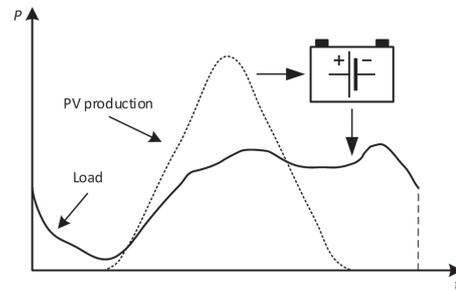


Fig. 10. Photovoltaic with a battery storage

3.5. Voltage and frequency regulation

In the power system it is required to maintain voltage and frequency to keep the power system stability. Voltage is regulated with the reactive power and frequency is regulated with the active power. Voltage and frequency changes in power systems with integrated renewables are more significant and occur more often so there is a need for advanced voltage and frequency regulation systems [37].

Grid voltage should be maintained within defined value range at consumption level. Grid voltages, out of range, may cause some electrical devices malfunction or even damage [37]. If voltage is lower than the nominal value, injection of the reactive power is required and if voltage is higher than the nominal value, absorption of the reactive power is required. In conventional power systems voltage regulation is done by static or synchronous compensation which injects reactive power in the grid when needed. There is ability to install the battery storage which can absorb reactive power to decrease voltage when voltage is high and inject the additional reactive power to increase voltage when voltage is low.

Frequency in grid should be maintained within defined value range and it can be done with balance between load and power generation [38]. When load power is higher than generated power, frequency is lower than the nominal value and injection of the additional active power in grid is required. When generated power is higher than load power, frequency is higher than the nominal value and decrease of generated power or increase of load power is required. In conventional power systems frequency regulation is done by increasing or decreasing generator output power. There is an ability to install a battery storage which can absorb power excess from the grid to decrease frequency when it is too high and inject power to the grid to increase frequency when it is too low.

3.6. Uninterruptible power supply

Uninterruptible power supplies (UPS) have a task to supply electrical and electronic equipment with constant power during power outages or state of emergency. UPS are usually used in the critical applications where the main power supply interruption, even for a

short time, can have serious, or even dangerous, consequences [39]. Critical applications are: control and monitoring systems in substations, power plants and industrial facilities, emergency lighting, medical equipment, computer systems, data centers, telecommunication systems and base stations, airports, lighting and signalization in tunnels, etc.

Main parts of a typical UPS system are battery bank, rectifier, inverter and static switch [40]. Battery banks have a task to store electrical energy which will be used during power outages, lead-acid technology was usually used and now, lithium-ion technology is often in use. Rectifiers are used to connect AC grids with DC batteries and enable to charging batteries with energy from the grid. Inverter is used to connect DC batteries with AC load and enable to supply load with energy from batteries. Static switch has a task to bypass converters when the rectifier or inverter are not working and enable to power load with energy directly from the grid.

4. BATTERY TECHNOLOGY COMPARISON

Comparison of overviewed battery technologies is done according to technical characteristics given in table 1 – 6. Comparison of characteristics is given through following six diagrams which are uniformed for better technology comparison. The first four diagrams are two-dimensional, one technology characteristic is shown on x-axis while another is showed on y-axis. The last two diagrams are one-dimensional, battery technologies are shown on x-axis while technology characteristic is shown on y-axis.

Diagram which shows specific power and specific energy for different battery storage technologies is presented in Fig. 11. Lead-acid, nickel-cadmium, nickel-metal hydride and vanadium-redox flow battery have the low specific power and specific energy. Sodium-sulfur technology has the low specific power, but the high specific energy, from 150 to 240 Wh/kg. For the lithium-ion battery technology, specific energy is from 80 to 250 Wh/kg and specific power is from 200 to 2000 W/kg, which is more than other technologies.

Diagram which shows power density and energy density for different battery storage technologies is presented in Fig. 12. Vanadium-redox flow battery has very low power density, from 2.5 to 3 kW/m³ and very low energy density, from 10 to 33 kWh/m³. It is because a large space needed for two electrolyte tanks placement. For li-ion technology, energy density is from 95 to 500 kWh/m³ and power density is from 50 to 800 kW/m³, which is more than other technologies.

Diagram which shows power cost and energy cost for different battery storage technologies is presented in Fig. 13. Most expensive battery technology is lithium-ion with power cost from 1000 to 3400 €/kW and energy cost from 500 to 2100 €/kWh. Battery technology with the lowest price is lead-acid with power cost from 250 to 500 €/kW and energy cost from 40 to 170 €/kWh.

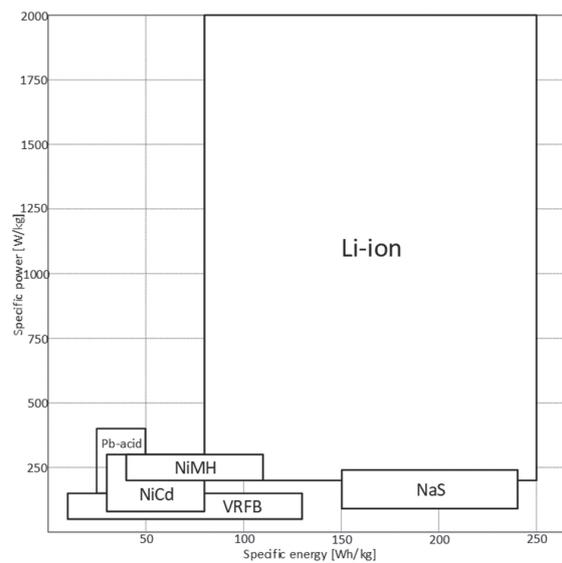


Fig. 11. Specific power to specific energy

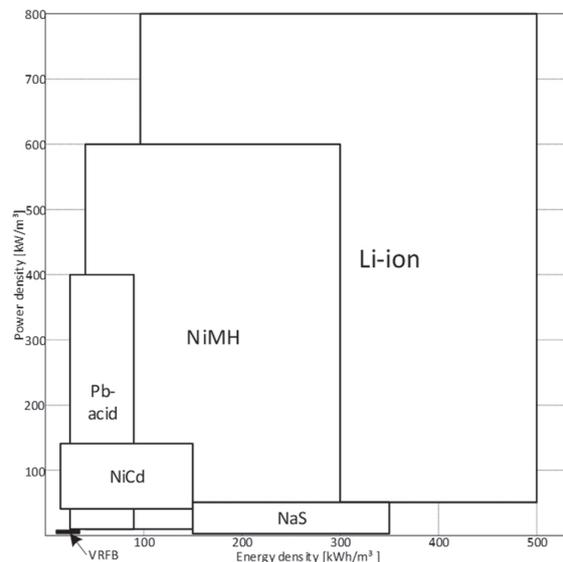


Fig. 12. Power density to energy density

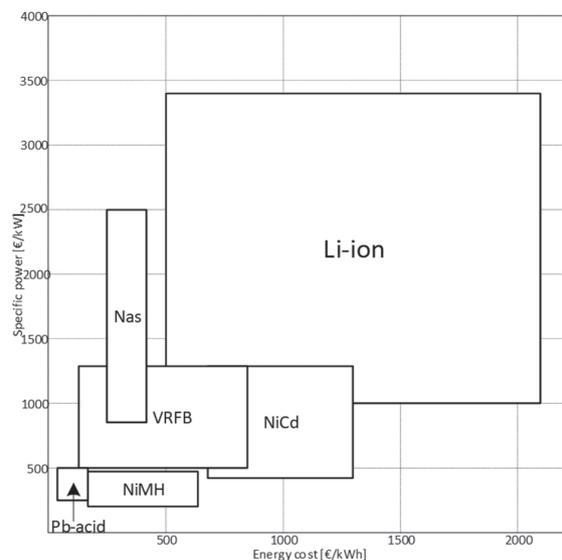


Fig. 13. Cost of battery technologies

A diagram which shows the lifetime of battery technologies is presented in Fig. 14. Lifetime for the lithium-ion is from 5 to 15 years and the number of lifetime cycles is from 1000 to 10000. Nickel-cadmium battery technology has the longest lifetime, from 10 to 20 years. Sodium-sulfur battery technology has the greatest number of lifetime cycles, from 2500 to 40000.

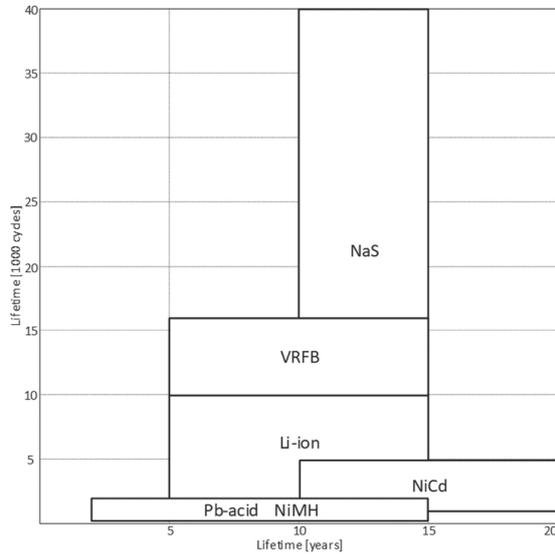


Fig. 14. Lifetime of battery technologies

Fig. 15 shows cell voltage for different battery technologies. Nickel-cadmium, nickel-metal hydride and vanadium-redox flow battery have lowest cell voltage. Sodium-sulfur and lead-acid have some higher cell voltage, Highest cell voltage has the lithium-ion technology, from 2.5 to 5 V.

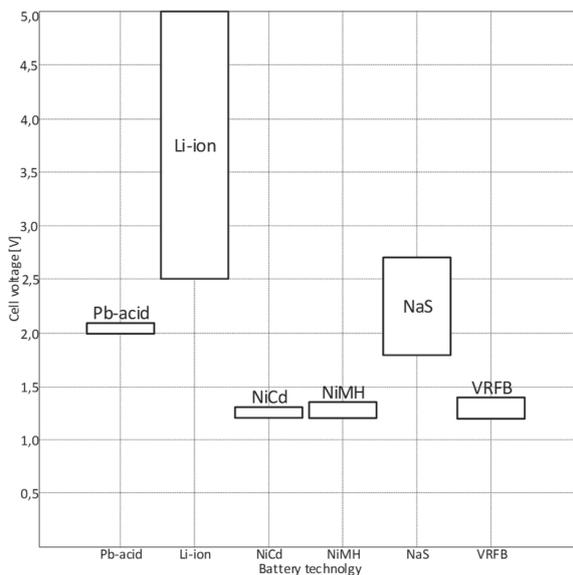


Fig. 15. Cell voltage of battery technologies

Fig. 16 shows efficiency of different battery technologies. Nickel-metal hydride has lowest efficiency, from 50 to 80 %. Highest efficiency is of the lithium-ion technology, from 75 to 97 %.

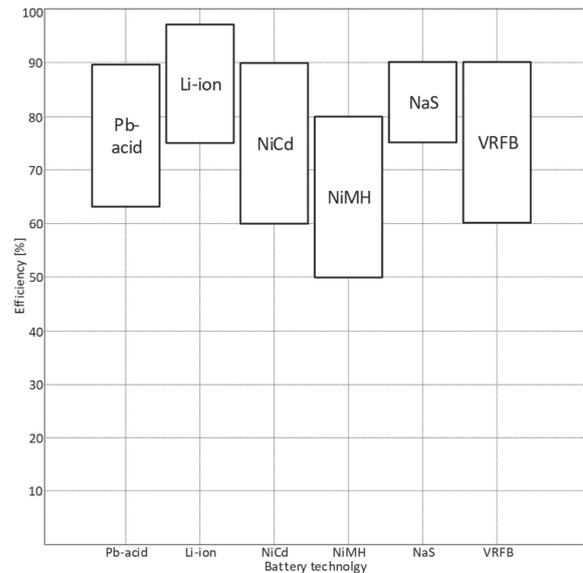


Fig. 16. Efficiency of battery technologies

According to diagrams in Fig. 11 – 16, the overall table of compared technology characteristics is given. Table shows overall comparison between overviewed battery technologies according to most important technical characteristics. Through diagrams, ten battery technical characteristics are compared as follows: specific energy and specific power, energy density and power density, energy cost and power cost, lifetime and lifetime cycles, cell voltage and the last, efficiency. Values in the overall table are shown in grayscales. The best technology characteristics are shown in a darker shade and the worst technology characteristics are shown in a lighter shade.

5. CONCLUSIONS

In this paper, six battery technologies, that are most often used for grid applications, are overviewed. Main characteristics of the lead-acid, lithium-ion, nickel-cadmium, nickel-metal hydride, sodium-sulfur and vanadium-redox flow batteries are presented and their characteristics are compared. It cannot be concluded that one technology is better than others. Some battery technologies are suitable for different applications while others have a specific application in which they have proved very good. All battery technologies are good for implementation in moderate climate conditions but only few technologies can withstand implementation in extreme climate conditions.

Lead-acid is in use for a long time and the main advantage of this technology is the very low price. Lead-acid technology is suitable for different stationary applications because of good efficiency and the high cell voltage. But lead-acid battery technology has the low number of lifetime cycles. Lithium-ion is now the most advanced and the most used battery technology. According to the overall table, lithium-ion battery technology is the most suitable for different applications because of highest specific power and energy, highest

Table 7. Comparison of battery technologies

Characteristics	Pb-acid	Li-ion	NiCd	NiMH	NaS	VRFB
Specific energy [Wh/kg]	25 – 50	80 – 250	30 – 80	40 – 110	150 – 240	10 – 130
Specific power [W/kg]	150 – 400	200 – 2000	80 – 300	200 – 300	90 – 230	50 – 150
Energy density [kWh/m ³]	25 – 90	95 – 500	15 – 150	40 – 300	150 – 350	10 – 33
Power density [kW/m ³]	10 – 400	50 – 800	40 – 140	10 – 600	1.2 – 50	2.5 – 33
Energy cost [€/kWh]	40 – 170	500 – 2100	680 – 1300	170 – 640	250 – 420	130 – 850
Power cost [€/kW]	250 – 500	1000 – 3400	420 – 1300	200 – 470	850 – 2500	500 – 1300
Lifetime [years]	2 – 15	5 – 15	10 – 20	2 – 15	10 – 15	5 – 15
Lifetime cycles [cycles]	250 – 2000	100 – 10000	1000 – 5000	300 – 1800	2500 – 40000	10000 – 16000
Cell voltage [V]	2 – 2.1	2.5 – 5	1.2 – 1.3	1.2 – 1.35	1.8 – 2.71	1.2 – 1.4
Efficiency [%]	63 – 90	75 – 97	60 – 90	50 – 80	75 – 90	75 – 90

power and energy density, highest cell voltage and highest efficiency. But lithium-ion battery technology has the highest price in comparison to other technologies.

Nickel-cadmium is good because of the ability of operation in extreme conditions with very low and high temperatures, but because of bad environmental impact, nickel-cadmium is replaced with the nickel-metal hydride technology. For extremely cold and hot working conditions, the most suitable technology is nickel-metal hydride. It has relatively high power and energy density and low price, but with some lower efficiency.

For applications that require a high number of lifetime cycles, the most suitable technology is sodium-sulfur because of the high number of lifetime cycles. Sodium-sulfur battery technology has high specific energy and energy density, high cell voltage and good efficiency. Main disadvantage of sodium-sulfur technology is the high working temperature.

Vanadium-redox flow batteries are future storage technology with the ability of long-time energy storage, VRFB technology has the high number of lifetime cycles. Main disadvantages of this technology are very low energy and power density and large space required for battery placement.

6. ACKNOWLEDGEMENT

This work was funded by the European Union through the European Regional Development Fund's Operational Programme Competitiveness and Cohesion under project KK.01.1.1.04.0034 Connected Stationary Battery Energy Storage.

7. REFERENCES

- [1] X. Zhou, Y. Lin, Y. Ma, "The Overview of Energy Storage Technology", Proceedings of the International Conference on Mechatronics and Automation, Beijing, China, 2-5 August 2015, pp. 43-48.
- [2] I. Hadjipaschalis, A. Poullikkas, V. Efthimiou, "Overview of current and future energy storage technologies for electric power applications", Renewable and Sustainable Energy Reviews, Vol. 13, No. 6-7, 2009, pp. 1513-1522.
- [3] L. Branders, D. Sprake, Y. Vagapov, H. Tun, "Analysis of Electrical Energy Storage Technologies for Future Electric Grid", Proceedings of the IEEE NW Russia Young Researchers in Electrical and Electronic Engineering Conference, St. Petersburg, Russia, 2-3 February 2016, pp. 513-518.

- [4] P. Nikolaidis, A. Poullikkas, "A comparative review of electrical energy storage systems for better sustainability", *Journal of Power Technologies*, Vol. 97, No. 3, 2017, pp. 220-245.
- [5] S. K. Fayegh, M.A. Rosen, "A review of energy storage types, applications and recent developments", *Journal of Energy Storage*, Vol. 27, 2020.
- [6] A. Saez-de-Ibarra, A. Milo, H. Gaztañaga, I. Etxeberria-Otadui, P. Rodriguez, S. Bacha, V. Debusschere, "Analysis and Comparison of Battery Energy Storage Technologies for Grid Applications", *Proceedings of the IEEE Grenoble Conference*, Grenoble, France, 16-20 June 2013.
- [7] X. Fan, B. Liu, J. Liu, J. Ding, X. Han, Y. Deng, X. Lv, Y. Xie, B. Chen, W. Hu, C. Zhong, "Battery Technologies for Grid Level Large Scale Electrical Energy", *Transaction of Tianjin University*, Vol. 26, No. 2, 2020, pp. 92-103.
- [8] M. Morris, S. Tosunoglu, "Comparison of rechargeable battery technologies", *Proceedings of the ASME Early Career Technical Conference*, Atlanta, Georgia, USA, 2-3 November 2012.
- [9] Y. Cho, H. A. Gabbar, "Review of energy storage technologies in harsh environment", *Safety in Extreme Environments*, Vol. 1, 2019, pp. 11-25.
- [10] E. O. Ogunniyi, H. Pienaar, "Overview of Battery Energy Storage System Advancement for Renewable (Photovoltaic) Energy Applications", *Proceedings of the International Conference on the Domestic Use of Energy*, Cape Town, South Africa, 4-5 April 2017.
- [11] N. Ertugul, "Battery Storage Technologies, Applications and Trend in Renewable Energy", *Proceedings of the IEEE International Conference on Sustainable Energy Technologies*, Hanoi, Vietnam, 14-16 November 2016, pp. 420-425.
- [12] G. J. May, A. Davidson, B. Monahov, "Lead batteries for utility energy storage: A review", *Journal of Energy Storage*, Vol. 15, 2018, pp. 145-157.
- [13] H. C. Hesse, M. Schimpe, D. Kucevic, A. Jossen, "Lithium-Ion Battery Storage for the Grid—A Review of Stationary Battery Storage System Design Tailored for Applications in Modern Power Grids", *Energies*, Vol. 10, No. 12, 2017.
- [14] T. Chen, Y. jin, H. Lv, A. Yang, M. Liu, B. Chen, Y. Xie, Q. Chen, "Applications of Lithium Ion Batteries in Grid Scale Energy Storage Systems", *Transactions of Tianjin University*, Vol. 26, No. 3, 2020, pp. 208-217.
- [15] Z. Xio, L. Du, X. Lv, Q. Wang, J. Huang, T. Fu, S. Li, "Evaluation and Analysis of Battery Technologies Applied to Grid Level Energy Storage Systems Based on Rough Set Theory", *Transactions of Tianjin University*, Vol. 26, 2020, pp. 228-235.
- [16] M. Sufyan, N. A. Rahim, M. M. Aman, C. K. Tan, S. R. Raihan, S. R. S. Raihan, "Sizing and applications of battery energy storage technologies in smart grid system: A review", *Journal of Renewable and Sustainable Energy*, Vol. 11, No. 1, 2019, p. 014105.
- [17] S. O. Masebinu, E. T. Akinlabi, E. Muzenda, A. O. Aboyade, C. Mbohwa, "A Review on Battery Technologies for Electrical Energy Storage", *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Rabat, Morocco, 11-13 April 2017, pp. 5711-5719.
- [18] L. Chang, W. Zhang, S. Xu, K. Spence, "Review on Distributed Energy Storage Systems for Utility Applications", *CPSS Transactions on Power Electronics and Applications*, Vol. 2, No. 4, 2017, pp. 267-276.
- [19] W. Xin, L. Yun, "Analysis of Energy Storage Technology and Their Application for Micro Grid", *Proceedings of the International Conference on Computer Technology, Electronics and Communication*, Sanya, China, 19-21 December 2017, pp. 972-975.
- [20] D. U. Sauer, M. Kleimaier, W. Glaunsinger, "Relevance of energy storage in future distribution networks with high penetration of renewable energy sources", *Proceedings of the 20th International Conference on Electricity Distribution*, Prague, 8-11 June 2009.
- [21] H. Keshan, J. Thornburg, T. S. Ustun, "Comparison of lead-acid and lithium ion batteries for stationary storage in off-grid energy systems", *Proceedings of the 4th Clean Energy and Technology Conference*, Kuala Lumpur, Malaysia, 14-15 November 2016, pp. 1-7.
- [22] D. Deng, "Li-ion batteries: basics, progress, and challenges", *Energy Science & Engineering*, Vol. 3, No. 5, 2015, pp. 385-418.

- [23] B. Scrosati, J. Garche, "Lithium batteries: Status, prospects and future", *Journal of Power Sources*, Vol. 195, No. 9, 2009, pp. 2419-2430.
- [24] R. Kasim, A. R. Abdullah, N. A. Selemat, M. S. Safwan, M. Basir, M. Z. Ramli, "Nickel-Cadmium battery analysis using spectogram", *ARPN Journal of Engineering and Applied Sciences*, Vol. 11, No. 6, 2016, pp. 3975-3979.
- [25] M. Zahran, A. Atef, "Electrical and Thermal Properties of NiCd Battery for Low Earth Orbit Satellite's Applications", *Proceedings of the 5th WSEAS International Conference on Power Systems*, Lisbon, Portugal, 22-24 September 2006, pp. 122-130.
- [26] K. H. Young, "Research in Nickel/Metal Hydride Batteries 2017", *Batteries*, Vol. 4, No. 9, 2018.
- [27] S. L. Lin, K. L. Huang, I. C. Wang, I. C. Chou, Y. M. Kuo, C. K. Hung, C. Lin, "Characterization of spent nickel-metal hydride batteries and a preliminary economic evaluation of the recovery processes", *Journal of the Air & Waste Management Association*, Vol. 66, No. 3, 2016, pp. 296-306.
- [28] J. Tarabay, N. Karami, "Nickel Metal Hydride Battery: Structure, Chemical Reaction, and Circuit Model", *Proceedings of the 3rd International Conference on Technological Advances in Electrical, Electronics and Computer Engineering*, Beirut, Lebanon, 29 April - 1 May 2015, pp. 22-26.
- [29] K. Saruta, "Long-term Performance of Sodium Sulfur Battery", *Proceedings of the IEEE 2nd Annual Southern Power Electronics Conference*, Auckland, New Zealand, 5-8 December 2016, pp. 1-5.
- [30] Z. Wen, "Study on Energy Storage Technology of Sodium Sulfur Battery and its Application in Power System", *Proceedings of the International Conference on Power System Technology*, Chongqing, China, 22-26 October 2006.
- [31] M. Vins, M. Sirovy, "Sodium-Sulfur Batteries for Energy Storage Applications", *Proceedings of the 20th International Scientific Conference on Electric Power Engineering*, Kouty nad Desnou, Czech Republic, Czech Republic, 15-17 May 2019.
- [32] O. Adeleke, A. Ukil, X. Zhang, "Vanadium Redox & Lithium Ion Based Multi-Battery Hybrid Energy Storage System for Microgrid", *Proceedings of the IEEE PES Asia-Pacific Power and Energy Engineering Conference*, Macao, 1-4 December 2019.
- [33] Y. R. Challapuram, G. M. Quintero, S. B. Bayne, A. S. Subburaj, M. A. Harral, "Electrical Equivalent Model of Vanadium Redox Flow Battery", *Proceedings of the IEEE Green Technologies Conference*, Lafayette, LA, USA, 3-6 April 2019.
- [34] R. Gupta, N. K. Sharma, P. Tiwari, A. Gupta, N. Nigam, A. Gupta, "Application of energy storage devices in power systems", *International Journal of Engineering, Science and Technology*, Vol. 2, No. 1, pp. 289-297, 2011.
- [35] M. Simeon, A. U. Adoghe, S. T. Wara, J. O. Olowoni, "Renewable Energy Integration Enhancement Using Energy Storage Technologies", *Proceedings of the IEEE PES/IAS PowerAfrica*, Cape Town, South Africa, 28-29 June 2018, pp. 864-868.
- [36] M. Miguel, T. Nogueira, F. Martins, "Energy storage for renewable energy integration: the case of Madeira Island, Portugal", *Proceedings of the 4th International Conference on Energy and Environment Research*, Porto, Portugal, 17-20 July 2017, pp. 251-257.
- [37] Y. Hua, X. Shentu, Q. Xie, Y. Ding, "Voltage/Frequency Deviations Control via Distributed Battery Energy Storage System Considering State of Charge", *Applied Sciences*, Vol. 9, No. 1148, 2019.
- [38] T. Tjark, M. Marten, J. Munderlein, H. Axelsen, S. Zurmühlen, M. Leuthold, D. U. Sauer, "Planning of Grid-Scale Battery Energy Storage Systems: Lessons Learned from a 5 MW Hybrid Battery Storage Project in Germany", *Proceedings of BattCon*, Orlando, Florida, USA, 12-14 May 2015, pp. 1-10.
- [39] J. O. Ushie, E. O. Ukem, B. E. Usibe, "Design, construction and testing of an uninterruptible power supply of 300 Watts capacity", *Global Journal of Engineering Research*, Vol. 11, No. 1, 2012, pp. 1-10.
- [40] D. R. Albu, F. V. Popentiu, "High reliability single-phase uninterruptible power supply", *Journal of Electrical and Electronics Engineering Research*, Vol. 3, No. 2, 2011, pp. 18-26.

INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

Published by Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia.

About this Journal

The International Journal of Electrical and Computer Engineering Systems publishes original research in the form of full papers, case studies, reviews and surveys. It covers theory and application of electrical and computer engineering, synergy of computer systems and computational methods with electrical and electronic systems, as well as interdisciplinary research.

Topics of interest include, but are not limited to:

- Power systems
- Renewable electricity production
- Power electronics
- Electrical drives
- Industrial electronics
- Communication systems
- Advanced modulation techniques
- RFID devices and systems
- Signal and data processing
- Image processing
- Multimedia systems
- Microelectronics
- Instrumentation and measurement
- Control systems
- Robotics
- Modeling and simulation
- Modern computer architectures
- Computer networks
- Embedded systems
- High-performance computing
- Parallel and distributed computer systems
- Human-computer systems
- Intelligent systems
- Multi-agent and holonic systems
- Real-time systems
- Software engineering
- Internet and web applications and systems
- Applications of computer systems in engineering and related disciplines
- Mathematical models of engineering systems
- Engineering management
- Engineering education

Paper Submission

Authors are invited to submit original, unpublished research papers that are not being considered by another journal or any other publisher. Manuscripts must be submitted in doc, docx, rtf or pdf format, and limited to 30 one-column double-spaced pages. All figures and tables must be cited and placed in the body of the paper. Provide contact information of all authors and designate the corresponding author who should submit the manuscript to <https://ijeces.ferit.hr>. The corresponding author is responsible for ensuring that the article's publication has been approved by all coauthors and by the institutions of the authors if required. All enquiries concerning the publication of accepted papers should be sent to ijeces@ferit.hr.

The following information should be included in the submission:

- paper title;
- full name of each author;
- full institutional mailing addresses;
- e-mail addresses of each author;
- abstract (should be self-contained and not exceed 150 words). Introduction should have no subheadings;
- manuscript should contain one to five alphabetically ordered keywords;
- all abbreviations used in the manuscript should be explained by first appearance;
- all acknowledgments should be included at the end of the paper;
- authors are responsible for ensuring that the information in each reference is complete and accurate. All references must be numbered consecutively and citations of references in text should be identified using numbers in square brackets. All references should be cited within the text;
- each figure should be integrated in the text and cited in a consecutive order. Upon acceptance of the paper, each figure should be of high quality in one of the following formats: EPS, WMF, BMP and TIFF;
- corrected proofs must be returned to the publisher within 7 days of receipt.

Peer Review

All manuscripts are subject to peer review and must meet academic standards. Submissions will be first considered by an editor-

in-chief and if not rejected right away, then they will be reviewed by anonymous reviewers. The submitting author will be asked to provide the names of 5 proposed reviewers including their e-mail addresses. The proposed reviewers should be in the research field of the manuscript. They should not be affiliated to the same institution of the manuscript author(s) and should not have had any collaboration with any of the authors during the last 3 years.

Author Benefits

The corresponding author will be provided with a .pdf file of the article or alternatively one hardcopy of the journal free of charge.

Units of Measurement

Units of measurement should be presented simply and concisely using System International (SI) units.

Bibliographic Information

Commenced in 2010.
ISSN: 1847-6996
e-ISSN: 1847-7003

Published: quarterly

Copyright

Authors of the International Journal of Electrical and Computer Engineering Systems must transfer copyright to the publisher in written form.

Subscription Information

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.

Postal Address

Faculty of Electrical Engineering,
Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b
31000 Osijek, Croatia

IJECES Copyright Transfer Form

(Please, read this carefully)

This form is intended for all accepted material submitted to the IJECES journal and must accompany any such material before publication.

TITLE OF ARTICLE (hereinafter referred to as "the Work"):

COMPLETE LIST OF AUTHORS:

The undersigned hereby assigns to the IJECES all rights under copyright that may exist in and to the above Work, and any revised or expanded works submitted to the IJECES by the undersigned based on the Work. The undersigned hereby warrants that the Work is original and that he/she is the author of the complete Work and all incorporated parts of the Work. Otherwise he/she warrants that necessary permissions have been obtained for those parts of works originating from other authors or publishers.

Authors retain all proprietary rights in any process or procedure described in the Work. Authors may reproduce or authorize others to reproduce the Work or derivative works for the author's personal use or for company use, provided that the source and the IJECES copyright notice are indicated, the copies are not used in any way that implies IJECES endorsement of a product or service of any author, and the copies themselves are not offered for sale. In the case of a Work performed under a special government contract or grant, the IJECES recognizes that the government has royalty-free permission to reproduce all or portions of the Work, and to authorize others to do so, for official government purposes only, if the contract/grant so requires. For all uses not covered previously, authors must ask for permission from the IJECES to reproduce or authorize the reproduction of the Work or material extracted from the Work. Although authors are permitted to re-use all or portions of the Work in other works, this excludes granting third-party requests for reprinting, republishing, or other types of re-use. The IJECES must handle all such third-party requests. The IJECES distributes its publication by various means and media. It also abstracts and may translate its publications, and articles contained therein, for inclusion in various collections, databases and other publications. The IJECES publisher requires that the consent of the first-named author be sought as a condition to granting reprint or republication rights to others or for permitting use of a Work for promotion or marketing purposes. If you are employed and prepared the Work on a subject within the scope of your employment, the copyright in the Work belongs to your employer as a work-for-hire. In that case, the IJECES publisher assumes that when you sign this Form, you are authorized to do so by your employer and that your employer has consented to the transfer of copyright, to the representation and warranty of publication rights, and to all other terms and conditions of this Form. If such authorization and consent has not been given to you, an authorized representative of your employer should sign this Form as the Author.

Authors of IJECES journal articles and other material must ensure that their Work meets originality, authorship, author responsibilities and author misconduct requirements. It is the responsibility of the authors, not the IJECES publisher, to determine whether disclosure of their material requires the prior consent of other parties and, if so, to obtain it.

- The undersigned represents that he/she has the authority to make and execute this assignment.
- For jointly authored Works, all joint authors should sign, or one of the authors should sign as authorized agent for the others.
- The undersigned agrees to indemnify and hold harmless the IJECES publisher from any damage or expense that may arise in the event of a breach of any of the warranties set forth above.

Author/Authorized Agent

Date

CONTACT

International Journal of Electrical and Computer Engineering Systems (IJECES)
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Josip Juraj Strossmayer University of Osijek
Kneza Trpimira 2b
31000 Osijek, Croatia
Phone: +38531224600,
Fax: +38531224605,
e-mail: ijeces@ferit.hr