JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

4/2021

Performance Evaluation of MC-CDMA Systems with Single User Detection Technique using Kernel and Linear Adaptive Method

R. Fateh, A. Darif, and S. Saf — Paper 1

Phonetic Segmentation using a Wavelet-based Speech Cepstral Features and Sparse Representation Classifier

I. Al-Hassani, O. Al-Dakkak, and A. Assami — Paper 12

Research Concerning Electromagnetic Emissions from Residential On-grid PV Systems

K. Maniak, J. Sobolewski, T. Tomczyk, K. Spalt and J. Wroński — Paper 23

Impact of Structures on the Operation of Air Traffic Radiolocation and Radionavigation Surveillance Systems

M. J. Grzybkowski, D. Niewiadomski, and M. Mora — Paper 32

Accurate Location of Fiber Cable Fault with OTDR

K. Borzycki and P. Gajewski — Paper 42

Fusion of Depth and Thermal Imaging for People Detection

W. Gutfeter and A. Pacut — Paper 53

Has the Internet Saved the Economy? Modeling Impact of ICT Sector and COVID-19 on GDP

M. Olender-Skorek, M. Sylwestrzak, and M. Szałański — Paper 61

Editor-in-Chief

Adrian Kliks, *Poznan University of Technology, Poland*

Steering Editor

Paweł Pławiak, *National Institute of Telecommunications, Poland*

Editorial Advisory Board

Hovik Baghdasaryan, *National Polytechnic University of Armenia, Armenia*
Naveen Chilamkurti, *LaTrobe University, Australia*
Luis M. Correia, *Instituto Superior Técnico, Universidade de Lisboa, Portugal*
Luca De Nardis, *DIET Department, University of Rome La Sapienza, Italy*
Nikolaos Dimitriou, *NCSR "Demokritos", Greece*
Ciprian Dobre, *Politechnic University of Bucharest, Romania*
Filip Idzikowski, *Poznan University of Technology, Poland*
Andrzej Jajszczyk, *AGH University of Science and Technology, Poland*
Albert Levi, *Sabanci University, Turkey*
Marian Marciniak, *National Institute of Telecommunications, Poland*
George Mastorakis, *Technological Educational Institute of Crete, Greece*
Constandinos Mavromoustakis, *University of Nicosia, Cyprus*
Klaus Mößner, *Technische Universität Chemnitz, Germany*
Imran Muhammad, *King Saud University, Saudi Arabia*
Mjumo Mzyece, *University of the Witwatersrand, South Africa*
Daniel Negru, *University of Bordeaux, France*
Ewa Orłowska, *National Institute of Telecommunications, Poland*
Jordi Perez-Romero, *UPC, Spain*
Michał Pióro, *Warsaw University of Technology, Poland*
Konstantinos Psannis, *University of Macedonia, Greece*
Salvatore Signorello, *University of Lisboa, Portugal*
Adam Wolisz, *Technische Universität Berlin, Germany*
Tadeusz A. Wysocki, *University of Nebrasca, USA*

Publications Staff

Content Editor: **Robert Magdziak**
Managing Editor: **Ewa Kapuściarek**
Technical Editor: **Julia Malińska**

on-line: ISSN 1899-8852

© Copyright by National Institute of Telecommunications, Warsaw 2021

Paper

Performance Evaluation of MC-CDMA Systems with Single User Detection Technique using Kernel and Linear Adaptive Method

Rachid Fateh, Anouar Darif, and Said Safi

Sultan Moulay Slimane University, Beni Mellal, Morocco

https://doi.org/10.26636/jtit.2021.151621

Abstract—Among all the techniques combining multi-carrier modulation and spread spectrum, the multi-carrier code division multiple access (MC-CDMA) system is by far the most widely studied. In this paper, we present the performance of the MC-CDMA system associated with key single-user detection techniques. We are interested in problems related to identification and equalization of mobile radio channels, using the kernel method in Hilbert space with a reproducing kernel, and a linear adaptive algorithm, for MC-CDMA systems. In this context, we tested the efficiency of these algorithms, considering practical frequency selective fading channels, called broadband radio access network (BRAN), standardized for MC-CDMA systems. As far as the equalization problem encountered after channel identification is concerned, we use the orthogonality restoration combination (ORC) and the minimum mean square error (MMSE) equalizer techniques to correct the distortion of the channel. Simulation results demonstrate that the kernel algorithm is efficient for practical channels.

Keywords—BRAN channels, equalization, kernel method, MC-CDMA, reproducing kernel Hilbert space.

1. Introduction

The problem of identifying a finite impulse response (FIR) system is the subject of interest for a large number of researchers [1]–[7]. Indeed, identifying a system means finding a set of parameters that form a mathematical model linking the different variables of the system [8]. Identification of linear systems has been carried out for decades using stochastic gradient algorithms. However, today, because of their complexity, the systems are becoming more non-linear. Given the increasing use of non-linear models in real systems, numerous resolution methods for the identification of non-linear systems, such as Volterra filters, neural networks and kernel methods, have been developed [9]–[14]. Kernel-based methods have been highly successful success in a wide range of fields over the past decade [15]. They are founded on the robust mathematical framework of reproducing kernel Hilbert spaces (RKHS), creating an interesting framework for the development of adaptive non-linear filters [8], [16]. Concerns related to the application of these algorithms have led us to consider the problem of equalization of MC-CDMA systems using BRAN channels. In order to solve these issues, several authors have suggested additional approaches in [17]–[24]. Innovative multi-carrier CDMA [25] transmission techniques are emerging as high-potential solutions for the fourth-generation of cellular networks, due to the many advantages of MC-CDMA, such as reduced inter-symbol interference (ISI), excellent spectral efficiency, as well as the ability to prevent frequency-selective fading, multi-carrier modulation and demodulation using inverse fast Fourier transform (IFFT) at the transmitter and fast Fourier transform (FFT) at the receiver, as these are less complicated to implement. MC-CDMA has acquired a great deal of importance in multi-user wireless communication systems [26]. In this paper, the main goal is to analyze the performance of MC-CDMA systems used for downlink over BRAN channels. We considered two practical frequency-selective fading channels called broadband radio access network (BRAN C and BRAN E), standardized by the European Telecommunications Standards Institute (ETSI) [27], [28]. Performance of the downlink MC-CDMA systems' bit error rate (BER) using the BRAN measures is shown and compared with the results achieved with the kernel LMS and LMS algorithms.

The present paper is arranged as follows. In Section 2, we introduced the Hammerstein system identification architecture. In Section 3, derivation of the least mean square algorithm is presented. Section 4 reviews some basic concepts of the kernel methods, and the kernel least mean squares identification algorithm is presented. Section 5 depicts the model of the MC-CDMA system. Monte Carlo simulations of those algorithms are presented in Section 6 and, finally, Section 7 concludes the paper.

2. System Architecture

Let us consider a Hammerstein model represented in Fig. 1. It consists of a non-linear static function followed by an invariant time linear block.

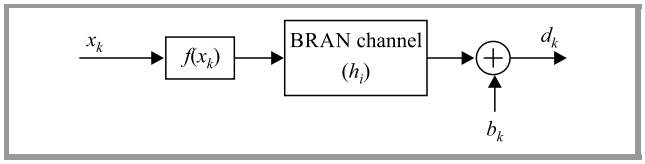


Fig. 1. Basic structure of a Hammerstein system with additive noise b_k [9].

The system's output for a defined input signal x_k, where $k = 0, 1, \ldots$, is obtained as follows:

$$d_k = \sum_{i=0}^{L-1} h_i f(x_{k-i}) + b_k, \tag{1}$$

where x_k is the symbol emitted by the source at moment k, the channel impulse response $\{h_i\}_{i=0,1,\ldots,L-1}$, L is the order of FIR system, and $f(.)$ is the non-linearity.
Throughout this paper, we rely on the following hypotheses concerning the system:

- input x_k, is independent and identically distributed (i.i.d.) with the zero mean.
- additive noise b_k is Gaussian and independent of x_k and d_k.
- non-linearity $f(.)$ is continuous and invertible for any finite x.
- there is no delay in the system, i.e. $h_0 \neq 0$.

3. LMS Algorithm

The least mean squares (LMS) [29] algorithm is one of the most popular algorithms for calculating coefficients of a finite impulse response filter. It is used to update the adaptive filter weights at each iteration:

$$\omega_{n+1} = \omega_n + \mu_{LMS} e_n \mathbf{x}_n, \tag{2}$$

where $\mathbf{x}_n = [x_n, x_{n-1}, \ldots, x_{n-L+1}]^\top$ is the reference signal vector of length L at moment n, $\omega_n = \omega_{0,n}, \omega_{1,n}, \ldots, \omega_{L-1,n}]^\top \in \mathbb{R}^L$ is the weight vector, $e_n = d_n - y_n$ is the error between the desired signal and the output filter and μ_{LMS} is known as the algorithm's step-size parameter. Its value has an impact on the performance of the LMS algorithm. In order to ensure the convergence of the weighting vector, it is necessary for the convergence step to be included in the interval below [30]:

$$0 < \mu_{LMS} < \frac{2}{\lambda_{max}}, \tag{3}$$

with λ_{max} being the maximum eigenvalue of the autocorrelation matrix, their values are non-negative.
The main disadvantage of this algorithm is the degradation of its performance when resolving non-linear problems.

4. Kernel-Based Adaptive Filters

Kernel methods undergo intense development these days, as they drive progress, both in terms of computational cost and performance achieved. They are based on a central principle known as the "kernel trick", exploited for the first time with the support vector machine (SVM) [31], adding a non-linear character to many originally linear methods. In practice, the kernel trick consists in rewriting an algorithm in which all relations between data inputs may be written as inner products, and replacing these products by scalar functions of two variables, where the original data input has been mapped to a non-linear Hilbert space (infinite-dimensional) using the Mercers theorem [13]:

$$\kappa(x_i, x_j) = \langle \psi(x_i), \psi(x_j) \rangle_{\mathcal{H}}, \quad \forall (x_i, x_j) \in \mathcal{E}^2, \tag{4}$$

where ψ maps $\mathcal{E}$ into a higher dimensions space $\mathcal{H}$ with an inner product $\langle ., . \rangle_{\mathcal{H}}$. Generally, $\dim(\mathcal{E}) \ll \dim(\mathcal{H})$.
In [13], [32]–[34], the authors have presented a detailed survey of the kernel method's characteristics. Figure 2 shows a functional diagram illustrating an adaptive filter based on the kernel, where $\{x_n\}_1^N$ is the vector of N most recent input signal samples, $\{y_n\}_1^N$ is the estimated desired response, and $\{e_n\}_1^N$ is the estimation error.

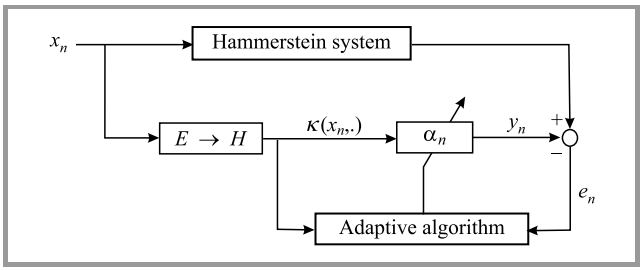


Fig. 2. Kernel-based adaptive system identification [35].

Any positively defined kernel can be seen as a scalar product in a functional space called the reproducing kernel Hilbert space (RKHS). This is therefore the property that a kernel must have to be validated. In order to properly determine the existence condition of a functional space $\mathcal{H}$, let us first start with some definitions [8], [36].
Definition 1 (positive definite kernel). A kernel is called positive definite if, for each input data point $\{x_i\}_{i=1}^N \in \mathcal{E}$, it satisfies the following condition:

$$\alpha^\top K \alpha = \sum_{i,j=1}^{N} \alpha_i \alpha_j \kappa(x_i, x_j) \geq 0, \tag{5}$$

for all $N \in \mathbb{N}$, $(x_1, \ldots, x_N) \in \mathcal{E}^N$ and $(\alpha_1, \ldots, \alpha_N) \in \mathbb{R}^N$.
We can assign an orthonormal base to a Hilbert space $\mathcal{H}$ in order to represent the elements of $\mathcal{H}$ from their coordinates.

The associated kernel must be a continuous, symmetric, normalized and positive definite function $\kappa : \mathcal{E} \times \mathcal{E} \rightarrow \mathbb{R}$. $\mathcal{E}$ is the input domain, a compact subset of $\mathbb{R}^N$.

Definition *2* (reproducing kernels and Hilbert spaces). Let $(\mathcal{H}, \langle .,. \rangle_{\mathcal{H}})$ represent a Hilbert space compromising functions of $\mathcal{E}$ in $\mathbb{R}$. The function $\kappa(x_i, x_j)$ of $\mathcal{E} \times \mathcal{E}$ in $\mathbb{R}$ is the reproducing kernel of $\mathcal{H}$, provided that $\mathcal{H}$ admits one, if there exists a function $\kappa(x,.) : x_i \longrightarrow \kappa(x, x_j)$ belongs to $\mathcal{H}$, for any $x \in \mathcal{E}$.

4.1. Kernel Least Mean Square

Here, we present the LMS kernel algorithm [37]–[39]. The basic idea is to run the linear LMS algorithm specified by Eq. (2) in the kernel feature space which is associated with the positive defined kernel κ. The sequence of samples is transformed by means of a feature map:

$$\begin{aligned} \psi \quad : \quad & \mathcal{E} \longrightarrow \mathcal{H} \\ & x \longrightarrow \kappa(x,.) \ . \end{aligned} \tag{6}$$

To construct the RKHS model, we use the Gaussian kernel:

$$\kappa(x_i, x_j) = \mathrm{e}^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}}, \quad \forall x_i, x_j \in \mathcal{E} \ , \tag{7}$$

where $\sigma > 0$ is the width of the kernel.

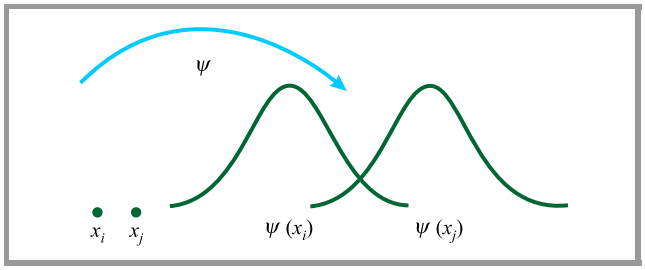


Fig. 3. Define a characteristic map.

Figure 3 shows data space $\mathcal{E}$ and space $\mathcal{H}$ induced by the kernel reproducing κ. This mapping represents each input point x by its similarity $\kappa(x,.)$ to all other points in the $\mathcal{E}$ domain. In order to build a feature space associated with ψ, the image of ψ must be transformed into a vector space with an inner product [32]. Now let us assume that the sequence of samples is transformed using feature map ψ. The LMS logic may be applied on the following transformed data:

$$\{(\psi(x_1), y_1), (\psi(x_2), y_2), \ldots (\psi(x_n), y_n), \ldots\}. \tag{8}$$

As a consequence of the linear structure of the cost function $J_\theta(n) = E[(d_n - y_n)^2]$ can be minimized compared to θ. We could solve this in the similar manner as in the LMS algorithm, using the instant-stochastic estimation, which gives:

$$\hat{\theta}_{n+1} = \hat{\theta}_n + \mu e_i \psi(x_i). \tag{9}$$

The main difference with LMS is that in Eq. (10) it is in a space of possibly infinite dimensional characteristics and its direct updating it would be virtually impossible. Instead, we will use each $\hat{\theta}_n$ of them to link to their initialization $\hat{\theta}_0$:

$$\hat{\theta}_n = \hat{\theta}_0 + \mu \sum_{i=1}^{N} e_i \psi(x_i). \tag{10}$$

By initializing the solution with a value 0, the solution can be expressed, after n iterations, as:

$$\hat{\theta}_n = \mu \sum_{i=1}^{N} e_i \psi(x_i). \tag{11}$$

By using the kernel trick, the following prediction function is obtained:

$$\begin{aligned} \langle \theta_n, \psi(x_n) \rangle_{\mathcal{H}} &= \mu \sum_{i=1}^{N} e_i \langle \psi(x_i), \psi(x_n) \rangle \\ &= \mu \sum_{i=1}^{N} e_i \kappa(x_i, x_n), \end{aligned} \tag{12}$$

where $\kappa(x_i, x_n)$ is a Mercer kernel, representing the inner product $\langle \psi(x_i), \psi(x_n) \rangle$ [39], and N is the number of training samples.

5. MC-CDMA System

The MC-CDMA technique is based on concatenation of the spread spectrum and multi-carrier modulation [42]–[40]. Instead of applying the spread-spectrum technique in the time domain, we apply it in the frequency domain, by modulating the different chips of the spreading code with OFDM subcarriers. More precisely, the complex symbol y_j specific to each user j is first multiplied by its user-specific Walsh-Hadamard spreading code:
$C_j = (c_{u,j})_{0<u\leq L_c}^{\mathsf{T}}$ of length L_c ,
where $(.)^{\mathsf{T}}$ designates the matrix transposition, then applied to the input of the multi-carrier modulator. Each sub-carrier transmits an information element multiplied by a chip of the specific code to this sub-carrier.
Figure 4 presents the general organization of a synchronous downlink based on the MC-CDMA technique in a scenario where length L_c of the spreading code is equal to number N_c of sub-carriers. In order to ensure orthogonality between subcarriers after the spreading function, the space between two adjacent subcarriers is proportional to the inverse of the T_c duration of an MC-CDMA symbol on each subcarrier. For the j-th transmitter, the emitted signal was represented as:

$$s_j(t) = \Re \left\{ \frac{1}{\sqrt{N_c}} \sum_{k=0}^{N_c-1} y_j c_{j,k} \mathrm{e}^{\mathrm{i}2\pi f_k t} \right\}, \text{ with } f_k = f_0 + \frac{k}{T_c} \ . \tag{13}$$

f_k designates the relative frequency, N_c corresponds to the number of subcarriers used, factor $\frac{1}{\sqrt{N_c}}$ is the power normalization term and $c_{j,k} \in \{-1, 1\}$ denotes the spreading code.

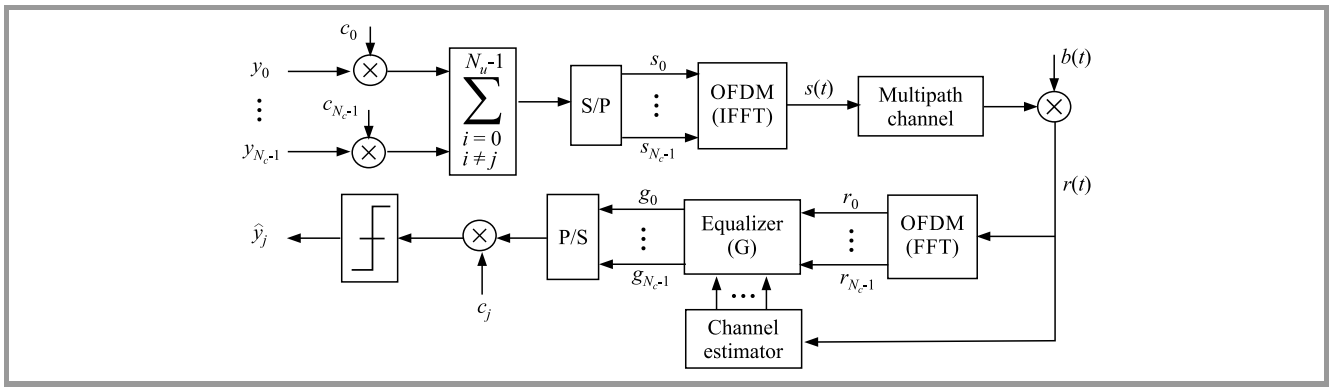


Fig. 4. MC-CDMA transmitter and receiver studied.

The propagation channel is time invariant and can be expressed in the delay domain by its impulse response:

$$h(\tau) = \sum_{l=0}^{L-1} \beta_l e^{j\theta_l} \delta(\tau - \tau_l) \ , \tag{14}$$

where β_l and θ_l are the magnitude and the phase of the channel, respectively.

The downlink MC-CDMA signal at the input of the receiver is noted $r(t)$ and is written as [43]:

$$r(t) = h(t) * s(t) + b(t) \ , \tag{15}$$

$$r(t) = \frac{1}{\sqrt{N_c}} \sum_{l=0}^{L-1} \sum_{j=0}^{N_u-1} \sum_{p=0}^{N_c-1} \Re\{\beta_l e^{j\theta} y_j c_{j,p} e^{2j\pi f_k(t-\tau_l)}\} + b(t) \ , \tag{16}$$

where N_u is the number of users, $\Re$ is a function of the real part, and $b(t)$ is additive white Gaussian noise.

The MC-CDMA symbol received may be expressed in a vector form by:

$$r = [r_0, r_1, ..., r_{N_c-1}] = HCy + b \ , \tag{17}$$

where r designates a vector made up of the values received on each subcarrier, and matrix H represents the channel coefficients diagonal matrix:

$$H = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ 0 & h_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_{N_c-1} \end{bmatrix} \ . \tag{18}$$

5.1. Single-user Equalization for the MC-CDMA System

The basic principle of equalization techniques is to reduce the effects of fading and interference, thus making it easier to make a decision about the received data symbols [17], [19]. Using the previous matrix notation, G, the diagonal matrix is composed of the g_j equalization coefficients and can be expressed by:

$$G = \begin{bmatrix} g_0 & 0 & \cdots & 0 \\ 0 & g_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & g_{N_c-1} \end{bmatrix} \ . \tag{19}$$

When equalized and de-spread according to sequence c_j of the user under consideration, estimation $\hat{y}_j$ of the emitted symbol may be expressed as:

$$\hat{y}_j = c_j.G.r \ . \tag{20}$$

Substituting Eq. (17) into Eq. (20) gives:

$$\hat{y}_j = c_j.G.H.C.y + c_j.G.b \ . \tag{21}$$

Matrix C represents N_c spreading codes c_j specific to each user j:

$$C = \begin{bmatrix} c_{0,0} & \cdots & c_{0,N_u-1} \\ & & \\ c_{L_c-1,0} & \cdots & c_{N_c-1,L_c-1} \end{bmatrix} \ , \tag{22}$$

thus:

$$\hat{y}_j = \sum_{p=0}^{N_c-1} C_{p,j}^2 g_p h_p y_j + \sum_{\substack{q=0 \\ q\neq j}}^{N_u-1} \sum_{p=0}^{N_c-1} C_{p,j} C_{p,q} g_p h_p y_q + \sum_{p=0}^{N_c-1} C_{p,j} g_p b_p \ . \tag{23}$$

Well-known single-user detection techniques include, inter alia, ORC and MMSE equalization.

5.1.1. Orthogonality Restoration Combination (ORC)

In order to completely cancel the phase and amplitude distortions provided by the channel, the ORC technique, also known as zero forcing (ZF), may be employed [23], [44]:

$$g_{orc} = \frac{1}{|h_p|}, \qquad \text{with } 0 \leq p \leq N_c - 1 \ . \tag{24}$$

Equation (23) allowing to obtain the estimation $\hat{y}_j$ of the symbol y_j of the user j is then written as:

$$\hat{y}_j = \sum_{p=0}^{N_c-1} C_{p,j}^2 y_j + \sum_{\substack{q=0 \\ q\neq j}}^{N_u-1} \sum_{p=0}^{N_c-1} C_{p,j} C_{p,q} y_q + \sum_{p=0}^{N_c-1} C_{p,j} \frac{1}{h_p} b_p \ . \tag{25}$$

Assuming that the spreading codes used at the transmitter level are orthogonal, we have:

$$\sum_{p=0}^{N_c-1} C_{p,j}C_{p,q} = 0, \quad \forall j \neq q. \tag{26}$$

The Eq. (25) reduces to:

$$\hat{y}_j = \sum_{p=0}^{N_c-1} C_{p,j}^2 y_j + \sum_{p=0}^{N_c-1} C_{p,j}\frac{1}{h_p}b_p.$$

However, the main drawback of this technique is that, for weak amplitudes of h_p, the multiplication by an inverse function of the channel results in a strong amplification of the noise, which rapidly increases the value of g_{orc}.

5.1.2. Minimum Mean Square Error (MMSE)

Equalization performed according to the MMSE criterion aims to minimize the mean square value of the error between vector S of the transmitted signal and its estimate $\hat{S} = GR$ generated at the output of the equalizer:

$$J = E\{|\varepsilon|^2\} = E\{|S - GR|^2\} \ , \tag{27}$$

where $R = (R_0, \ldots, R_k, \ldots, R_{N_c-1})$ and $G = (G_0, \ldots, G_k, \ldots, G_{N_c-1})$ are the symbol vector returned at the output of the FFT and the gain vector of the equalizer, respectively [45], [23]. This mean squared error J is minimal when the equalizer's G_k gains are such that both the received signal and the error signal are orthogonal, meaning that:

$$E\{\varepsilon R^{*\top}\} = 0 \ . \tag{28}$$

When the number of users is equal to length L_c of the code, the optimal equalizer coefficients, according to the mean square error minimization criterion, are:

$$g_{mmse} = \frac{h_p^*}{|h_p|^2 + \frac{1}{\gamma_p}}, \quad \text{with} \quad \gamma_p \neq 0 \ , \tag{29}$$

where $g_{mmse,p}$ is the p-th complex channel gain, the operation $*$ is the complex conjugate and γ_p is the signal-to-noise ratio for subcarrier p. The estimated received symbol, $\hat{y}_i$ of the symbol y_i of user i is represented by:

$$\hat{y}_j = \underbrace{\sum_{p=0}^{N_c-1} C_{p,j}^2 \frac{|h_p|^2}{|h_p|^2 + \frac{1}{\gamma_p}} y_j}_{\lambda} + \underbrace{\sum_{q=0}^{N_u-1}\sum_{p=0}^{N_c-1} C_{p,j}C_{p,q}\frac{|h_p|^2}{|h_p|^2+\frac{1}{\gamma_p}} y_p}_{\varsigma(q\neq j)} + \underbrace{\sum_{p=0}^{N_c-1} C_{p,j}\frac{h_p^*}{|h_p|^2 + \frac{1}{\gamma_p}} b_p}_{\eta} \ , \tag{30}$$

where λ is the usable signal portion, ς is the multiple access interference and η in the noise term.
If we assume that the code for spreading is orthogonal, that is:

$$\sum_{p=0}^{N_u-1} C_{p,j}C_{p,q} = 0, \quad \forall j \neq p \ . \tag{31}$$

So, Eq. (30) becomes:

$$\hat{y}_j = \sum_{p=0}^{N_u-1} C_{p,j}^2 \frac{|h_p|^2}{|h_p|^2 + \frac{1}{\gamma_p}} y_j + \sum_{p=0}^{N_u-1} C_{p,j}\frac{h_p^*}{|h_p|^2 + \frac{1}{\gamma_p}} b_p \ . \tag{32}$$

6. Simulation and Results

The simulation will allow us to study the performance of the adaptive kernel filtering algorithm. The mean square error (MSE) will be used to measure the accuracy of the estimated values as:

$$MSE(h,\hat{h}) = \frac{1}{p}\sum_{i=1}^{p}\left[\frac{h(i) - \hat{h}(i)}{h(i)}\right]^2 \ , \tag{33}$$

where $h(i)$ is the measured impulse response, $\hat{h}(i)$ is the estimated impulse response, and p represents the length of the impulse response.
MC-CDMA system's equalization performance in the single-user downlink scenario is evaluated using MMSE and ORC equalizers. This assessment is carried out using measured and estimated parameters of practical broadband radio access network (BRAN C and BRAN E) models to determine the bit error rate (BER), applying KLMS and LMS algorithms.

6.1. BRAN C and BRAN E Parameters

In Tables 1 and 2, we represent the measured values corresponding to BRAN C and BRAN E radio channels' impulse response. The impulse response of each model is:

$$h(n) = \sum_{i=1}^{p} M_i \delta(n - \tau_i) \ , \tag{34}$$

where $\delta(n)$ is the Dirac's function, τ_i is path's i time delay, M_i is path's i magnitude, where their impulse responses are between 0 and 1, $M_i \in N(0,1)$, $i = 1, \ldots, p$ and $p = 18$ is the number of paths.

Table 1
Delay and magnitudes of 18 targets of a BRAN C channel

Delay τ_i [ns]	**Magnitude** M_i [dB]	**Delay** τ_i [ns]	**Magnitude** M_i [dB]
0	-3.3	230	-3.0
10	-3.6	280	-4.4
20	-3.9	330	-5.9
30	-4.2	400	-5.3
50	0	490	-7.9
80	-0.9	600	-9.7
110	-1.7	730	-13.2
140	-2.6	880	-16.3
180	-1.5	1050	-21.2

Table 3 shows the various parameters of the MC-CDMA system implemented for the two BRAN channels: C and E.

Table 2
Delay and magnitudes of 18 targets of a BRAN E channel

Delay τ_i [ns]	**Magnitude** M_i [dB]	**Delay** τ_i [ns]	**Magnitude** M_i [dB]
0	-4.9	320	0
10	-5.1	430	-1.9
20	-5.2	560	-2.8
40	-0.8	710	-5.4
70	-1.3	880	-7.3
100	-1.9	1070	-10.6
140	-0.3	1280	-13.4
190	-1.2	1510	-17.4
240	-2.1	1760	-20.9

Table 3
Simulation parameters

Characteristic parameters	**Configuration**
Spreading code lengths L_c	64
Sampling frequency f_s	20 MHz
Number of sub-carriers N_c	64
Symbol times T_s	3 μs
Number of users N_u	64
Spreading codes	Walsh-Hadamard
OFDM modulation	FFT 64 samples
Channel model	BRAN (C and E)
Performance metrics	BER and MSE

6.2. Performance Results

6.2.1. ETSI BRAN C Channel

The parameters of the impulse response ETSI BRAN C radio channel estimated using the two algorithms are presented in Fig. 5, for a scenario with $SNR = 15$ dB, the number of input signal samples fixed at $N = 1024$ and 50 Monte Carlo iterations used. One may notice that the response estimated using the kernel algorithm matches the real values,

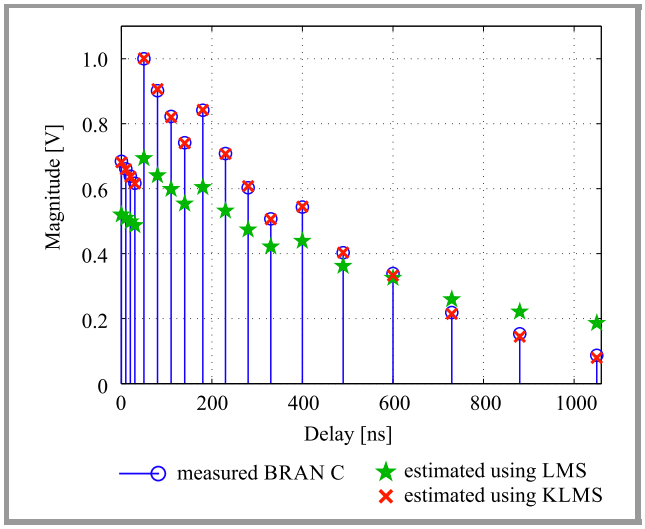


Fig. 5. Estimate of the BRAN C channel amplitude depending on the delay time.

but when estimating the channel impulse response using the LMS algorithm, there is an apparent difference between the estimated and the measured values.

The mean square error values for the two algorithms are presented in Fig. 6, for an SNR varying between 0 dB and 40 dB and the number of samples equaling $N = 1024$. We find that the kernel algorithm offers the best performance in terms of MSE criteria even in a highly noisy environment, compared to the LMS algorithm.

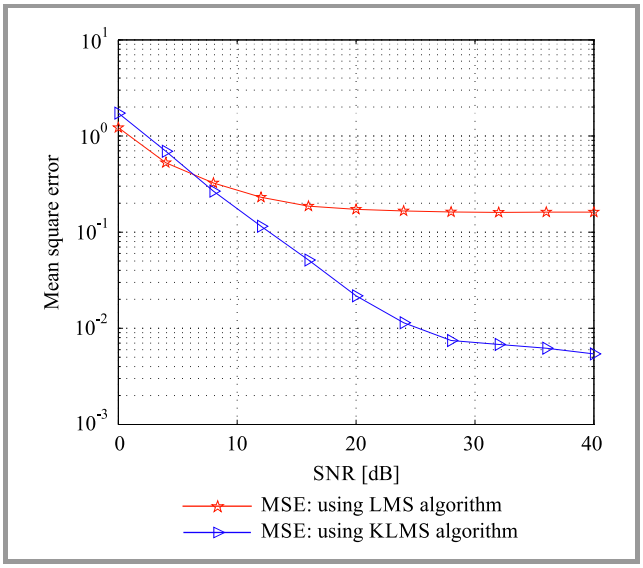


Fig. 6. Comparison of algorithms in terms of MSE for different *SNR* values and for the data length of $N = 1024$, BRAN C channel.

Figure 7 demonstrates estimates concerning of the amplitude and phase of the BRAN C channel, using the least mean square and the kernel least mean square algorithms, for a number of samples equaling $N = 1024$ and for an $SNR = 15$ dB. These results allow us to conclude that the kernel algorithm is more efficient compared with the LMS algorithm, because it allows us to have the same paces of the estimated (amplitude and phase) and those measured.

Figures 8 and 9 show, respectively, the BER simulation results, in the single-user and in the downlink scenario, for different SNRs obtained using ORC and MMSE equalizers, based on the measured parameters of the BRAN C channel and parameters estimated by the KLMS and LMS algorithms. From these simulation results (Fig. 8), one may conclude that the KLMS algorithm predicts the same performance values as those measured for the BRAN C channel, as opposed to the LMS algorithm which yields values quite different from their measured counterparts (BRAN C). Indeed, the Kernel LMS algorithm's BER values are inferior to those achieved by the LMS algorithm. From Figs. 8 and 9, we can conclude that:

- when using ORC equalization measured values. Both the two algorithms give the 1 bit error if we receive 10^2 bits for $SNR = 24$ dB, with an advantage over the kernel algorithm (Fig. 8),

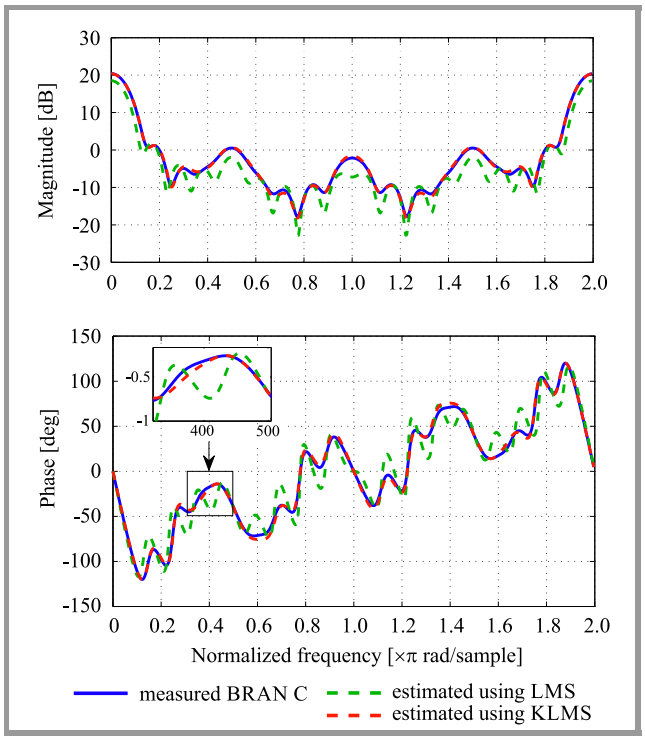


Fig. 7. Estimation of the amplitude and phase of the BRAN C channel, for the data length of $N = 1024$ and $SNR = 15$ dB.

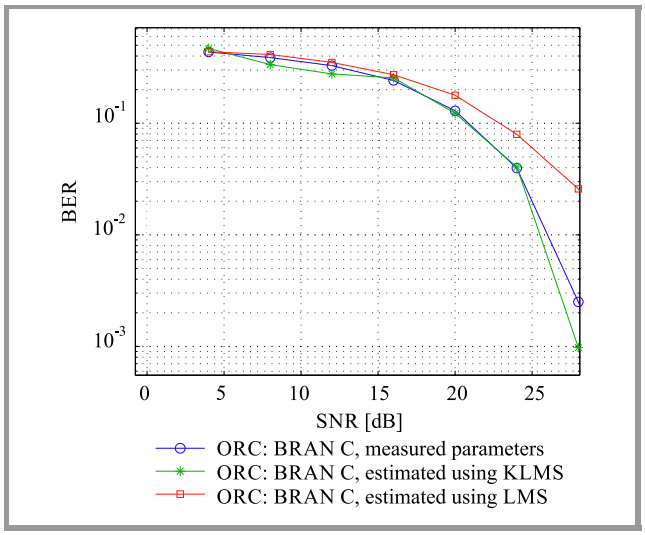


Fig. 8. BER vs. SNR of the estimated and measured BRAN C channel, using the ORC equalizer.

- in the case of the MMSE equalizer, if $SNR = 24$ dB we just get a 1 error bit, when we receive 10^4 bits, with an advantage over the kernel algorithm (Fig. 9).

We conclude that the MMSE technique is more efficient than the ORC technique in the case of the ETSI BRAN C radio channel.

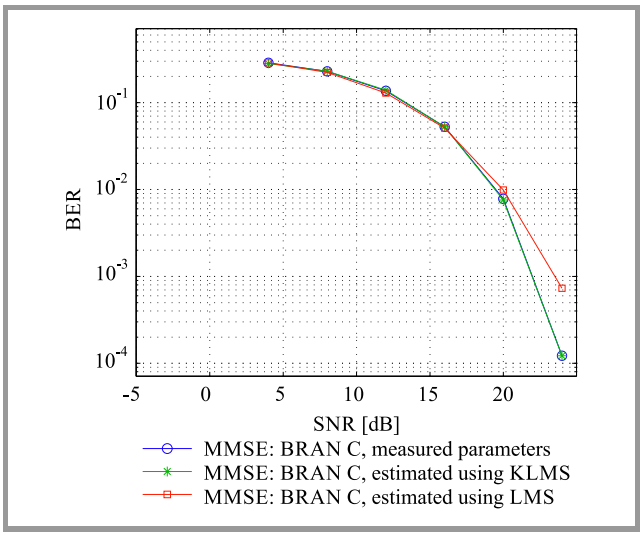


Fig. 9. BER vs. SNR of the estimated and measured BRAN C channel, using the MMSE equalizer.

6.2.2. ETSI BRAN E Channel

Figure 10 shows the estimated parameters of the BRAN E channel impulse response, as a function of the path delays, for a data length of $N = 1024$, $SNR = 15$ dB, and 50 Monte Carlo runs. One may notice that the best performance is obtained by the kernel LMS algorithm.

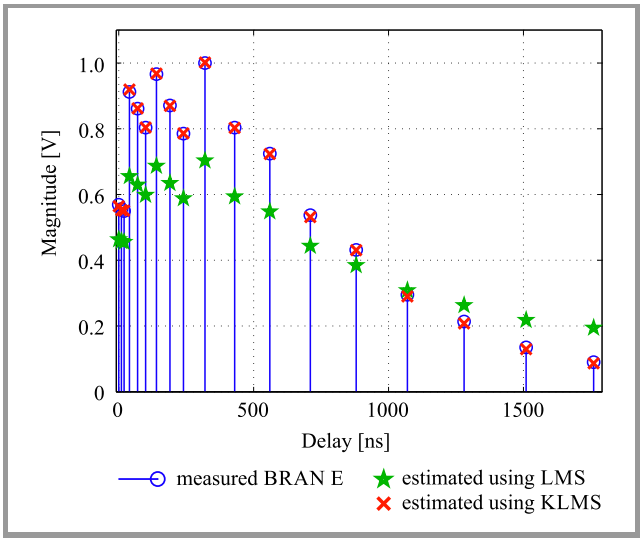


Fig. 10. BRAN E channel amplitude estimates depending on the time delay.

The mean square error values for the two algorithms are represented in Fig. 11, for different SNR and for the numbers of samples equaling $N = 1024$. One may notice that the performance of LMS is almost unaffected by any considerable disturbance, whereas the performance of the kernel has been significantly reduced due to its sensitivity to low SNR. The kernel algorithm achieves the best performance, with higher convergence speed and minimized MSE.

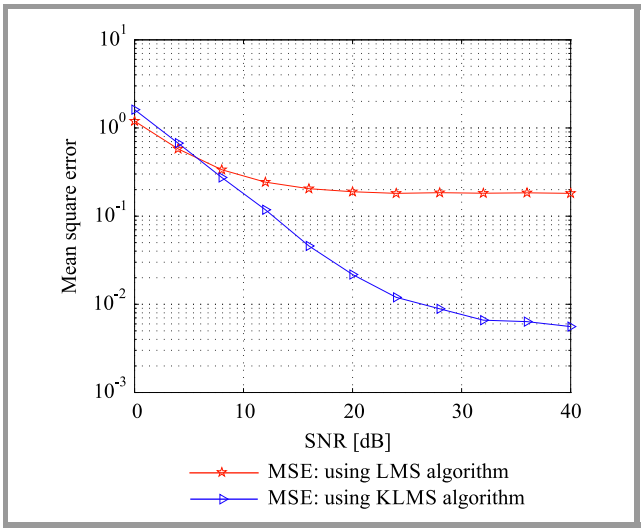


Fig. 11. Comparison of algorithms in terms of MSE for different *SNR* values and $N = 1024$, BRAN E channel.

As summarized in Figs. 11 and 6, it can be seen that the LMS algorithm is more effective in the case of the BRAN C channel, due to the greater level of fluctuations in the BRAN E channel. On the basis of the obtained results, one may notice that when the *SNR* value is lower than 5 dB, the value of MSE obtained with the KLMS algorithm (MSE $> 10^0$, if $SNR = 0$ dB) is slightly higher than the one obtained with LMS (MSE = 10^0, if $SNR = 0$ dB), but when $SNR > 5$ dB, we note the stability of the MSE value of the LMS so that no *SNR* value was affected, which testifies

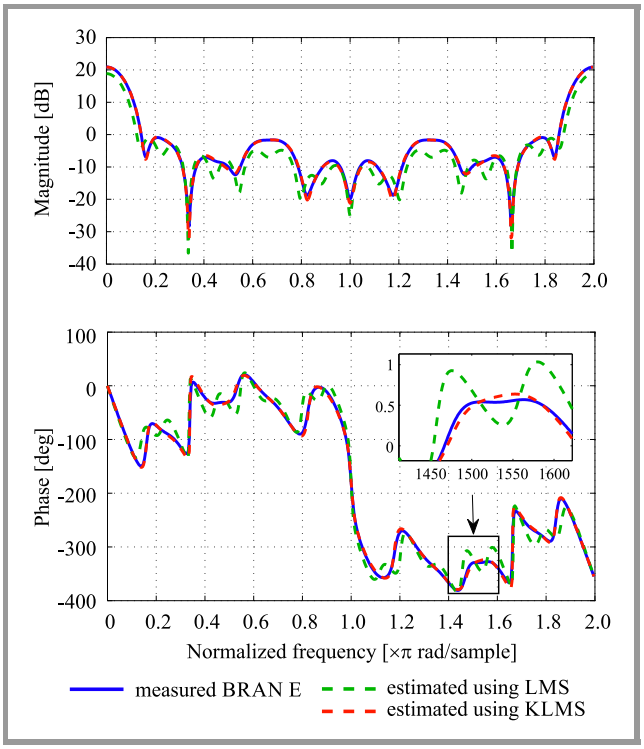


Fig. 12. Estimation of the amplitude and phase of the BRAN E channel, for a data length of $N = 1024$ and $SNR = 15$ dB.

to its incapacity to estimate the parameters. These results allow us to conclude that the kernel algorithm gives a good approximation of the model parameters to be identified. The estimation of the amplitude and phase of the BRAN E channel, using the two algorithms, is presented in Fig. 12, for a number of samples $N = 1024$ and $SNR = 15$ dB. The amplitude and phase estimated using the kernel algorithm have the same form as the measured data. Compared to the linear adaptive algorithm (LMS), we notice a difference between the shape of estimated amplitude and phase, and the shape of the actually measured parameters.

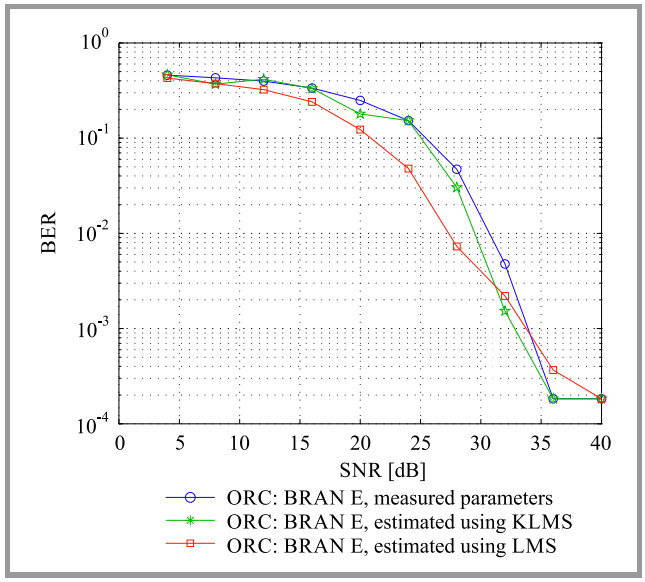


Fig. 13. BER vs. SNR of the estimated and measured BRAN E channel, using ORC equalizer.

The plot shown in Fig. 13 indicates BER for various SNRs, achieved using the kernel algorithm, and compares with the

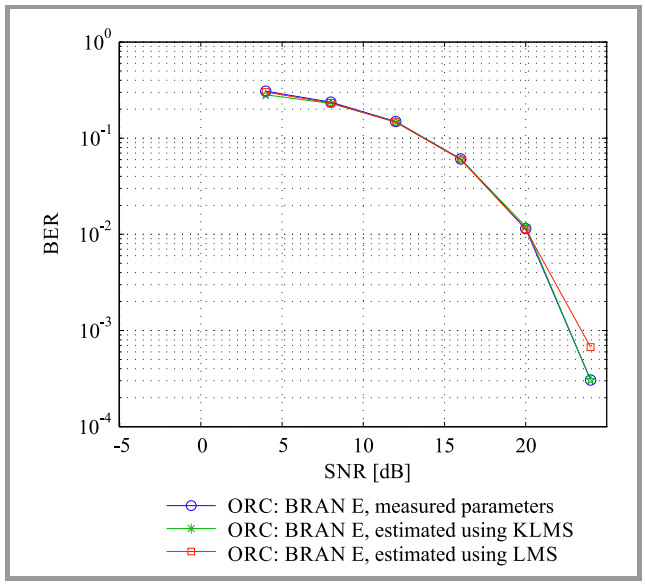


Fig. 14. BER vs. SNR of the estimated and measured BRAN E channel, using the MMSE equalizer.

values obtained with the LMS algorithm. Equalization is performed using the ORC equalizer.

Similarly, Fig. 14 presents the results of BER simulations for different SNRs obtained using the MMSE equalizer, based on the measured parameters of the BRAN E channel and the ones estimated with KLMS and LMS algorithms. We note that the efficiency of the KLMS algorithm is important compared to the LMS algorithm. According to Figs. 13 and 14, if $SNR = 36$ dB, the BER value is lower than 10^{-3} in the case of the ORC equalizer, but when using the MMSE technique, we obtain a BER value that is lower than 10^{-3} when the *SNR* equals 24 dB.

One may conclude that the MMSE equalizer offers the best performance in terms of BER for all channels studied (BRAN C and BRAN E).

7. Conclusion

In this paper, the performance of an MC-CDMA system in the downlink over BRAN channels was evaluated and analyzed, using the kernel LMS and LMS algorithms. These algorithms are used for estimating the parameters of the measured channels in different scenarios (BRAN C and BRAN E). The results presented in the identification part demonstrate that the kernel algorithm is effective and efficient for practical channels. As far as equalization of MC-CDMA systems is concerned, we obtained excellent bit error rate performance, mainly if the kernel last mean square algorithm was used. Future research includes the development of an extension of these algorithms to MIMO systems, new methods for identifying communication channels (i.e. using kernel methods) and development of new equalizers practically for MC-CDMA systems can be presented in the future.

References

[1] M. Pouliquen, T. Menard, E. Pigeon, O. Gehan, and A. Goudjil, "Recursive system identification algorithm using binary measurements", in *European Control Conference (ECC)*, Aalborg, Denmark, 2016, pp. 1353–1358 (DOI: 10.1109/ECC.2016.7810477).

[2] M. Zidane, S. Safi, M. Sabri, and A. Boumezzough, "Impulse response identification of minimum and non minimum phase channels", *Int. J. of Adv. Sci. and Technol.*, vol. 64, 2014, pp. 59–72 (DOI: 10.14257/ijast.2014.64.06).

[3] U. Soverini and T. Söderström, "Frequency domain identification of FIR models in the presence of additive input-output noise", *Automatica*, vol. 115, pp. 108879, 2020 (DOI: 10.1016/j.automatica.2020.108879).

[4] H. Zhang, T. Wang, and Y. Zhao, "Asymptotically efficient recursive identification of FIR systems with binary-valued observations", *IEEE Trans. on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 5, pp. 2687–2700, 2021 (DOI: 10.1109/TSMC.2019.2916022).

[5] U. Soverini and T. Söderström, "Frequency domain identification of ARX models in the presence of additive input-output noise", *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 6226–6231, 2017 (DOI: 10.1016/j.ifacol.2017.08.1023).

[6] R. Fateh and A. Darif, "Mean square convergence of reproducing kernel for channel identification: application to BRAN D channel impulse response", in *Int. Conf. on Business Intell.*, Springer: Cham, Switzerland, 2021, pp. 284–293 (DOI: 10.1007/978-3-030-76508-8_20).

[7] R. Fateh, A. Darif, and S. Safi, "Channel Identification of Non-linear Systems with Binary-Valued Output Observations Based on Positive Definite Kernels", in *E3S Web of Conferences, EDP Sciences*, vol. 297, pp. 01020, 2021 (DOI: 10.1051/e3sconf/202129701020).

[8] S. Van Vaerenbergh, "Kernel methods for nonlinear identification, equalization and separation of signals", Ph.D. dissertation, Universidad de Cantabria, Santander, Spain, 2010 [Online]. Available: https://gtas.unican.es/files/pub/tesis_Steven.pdf

[9] R. Castro Garcia, "Structured nonlinear system identification using kernel-based methods", Ph.D. dissertation, Faculty of Engineering Science, KU Leuven, Belgium, 2017 [Online]. Available: https://lirias.kuleuven.be/retrieve/473580

[10] S. Gupta, A. K. Sahoo, and U. K. Sahoo, "Volterra and Wiener model based temporally and spatio-temporally coupled nonlinear system identification: a synthesized review", *IETE Technical Review*, 2020, pp. 1–25 (DOI: 10.1080/02564602.2020.1732233).

[11] F. Pérez-Cruz and O. Bousquet, "Kernel methods and their potential use in signal processing", *IEEE Signal Process. Mag.*, vol. 21, no. 3, pp. 57–65, 2004 (DOI: 10.1109/MSP.2004.1296543).

[12] W. Liu, J. C. Principe, and S. Haykin, "Kernel adaptive filtering: a comprehensive introduction", vol. 57, *John Wiley & Sons*, 2011 (ISBN: 9780470447536).

[13] N. Aronszajn, "Theory of reproducing kernels", *Trans. of the American mathematical society*, vol. 68, no. 3, pp. 337–404, 1950 (DOI: 10.2307/1990404).

[14] S. Van Vaerenbergh and L. A. Azpicueta-Ruiz, "Kernel-based identification of Hammerstein systems for nonlinear acoustic echo-cancellation", in *IEEE Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, Florence, Italy, 2014. pp. 3739–3743 (DOI: 10.1109/ICASSP.2014.6854300).

[15] C. Richard, J. C. M. Bermudez, and P. Honeine, "Online Prediction of Time Series Data With Kernels", *IEEE Trans. on Signal Process.*, vol. 57, no. 3, pp. 1058–1067, 2009 (DOI: 10.1109/TSP.2008.2009895).

[16] O. A. Arqub and H. Rashaideh, "The RKHS method for numerical treatment for integrodifferential algebraic systems of temporal two-point BVPs", *Neural Computing and Applications*, vol. 30, no. 8, pp. 2595–2606, 2018 (DOI: 10.1007/s00521-017-2845-7).

[17] M. Boutalline, J. Antari, M. Gouskir, B. Bouikhalene, and S. Safi, "A robust approach for Proakis and BRAN channel identification and equalization of MC-CDMA systems", in *5th Int. Conf. on Optimization and App. (ICOA)*, Kenitra, Morocco, 2019, pp. 1–9 (DOI: 10.1109/ICOA.2019.8727665).

[18] S. Safi, M. Frikel, M. Pouliquen, I. Badi, Y. Khmou, and M. Boutalline, "MC-CDMA system identification and equalization using the LMS algorithm and Takagi-Sugeno fuzzy system", *IFAC Proc. Volumes*, vol. 46, no. 11, pp. 599–604, 2013 (DOI: 10.3182/20130703-3-FR-4038.00027).

[19] M. Zidane, S. Safi, M. Sabri, and A. Boumezzough, "Blind identification channel using higher order cumulants with application to equalization for MC-CDMA system", *Int. J. of Elec., Robot., Electron. and Commun. Engin.*, vol. 8, no. 2, pp. 369–375, 2014 (DOI: 10.5281/zenodo.1091694).

[20] M. S. Sakib and M. R. Ferdous, "Performance evaluation of MC-CDMA system for the next generation wireless network"', in *Int. Conf. on Innovation in Engin. and Technol. (ICIET)*, Dhaka, Bangladesh, 2018, pp. 1–5 (DOI: 10.1109/CIET.2018.8660783).

[21] M. Zidane, S. Safi, M. Sabri, A. Boumezzough, and M. Frikel, "Broadband radio access network channel identification and downlink MC-CDMA equalization", *Int. J. of Energy, Inform. and Commun.*, vol. 5, no. 2, pp. 13–34, 2014 (DOI: 10.14257/ijeic.2014.5.2.02).

[22] S. Safi, M. Frikel, M. M'Saad, and A. Zeroual, "Blind identification and equalization of downlink TCM coded MC-CDMA systems using cumulants", in *16th European Signal Process. Conf.*, IEEE. Lausanne, Switzerland, 2008, pp. 1–5, 2008 (ISSN: 22195491).

[23] M. Zidane, Mohammed, S. Safi, and M. Sabri, "Measured and estimated data of non-linear BRAN channels using HOS in 4G wireless communications", *Data in brief*, vol. 17, pp. 1136–1148, 2018 (DOI: 10.1016/j.dib.2018.02.005).

[24] M. Zidane, S. Safi, and M. Sabri, "Compensation of fading channels using partial combining equalizer in MC-CDMA systems", *J. of Telecommun. and Informat. Technol.*, no. 1, pp. 5–11, 2017 [Online]. Available: http://dlibra.itl.waw.pl/dlibra-webapp/Content/1962/ISSN_1509-4553_1_2017_5.pdf

[25] S. Chitra, and N. Kumaratharan, "Multimedia transmission in MC-CDMA using adaptive subcarrier power allocation and CFO compensation", *Int. J. of Electron.*, vol. 105, no. 2, pp. 289–302, 2018 (DOI: 10.1080/00207217.2017.1357201).

[26] S. Hara and R. Prasad, "Overview of multicarrier CDMA", *IEEE Commun. Mag.*, vol. 35, no. 12, pp. 126–133, 1997 (DOI: 10.1109/35.642841).

[27] "Broadband radio access networks (BRAN); High performance radio logical area network (HIPERLAN) Type 2; Requirements and architectures for wireless broadband access", *ETSI*, 1999 [Online]. Available: https://www.etsi.org/deliver/etsi_tr/101000_101099/101031/02.02.01_60/tr_101031v020201p.pdf

[28] "Broadband radio access networks (BRAN); HIPERLAN Type 2; Physical layer", *ETSI*, 2001 [Online]. Available: https://www.etsi.org/deliver/etsi_ts/101400_101499/101475/01.02.02_60/ts_101475v010202p.pdf

[29] R. Martinek, J. Rzidky, R. Jaros, P. Bilik, and M. Ladrova, "Least mean squares and recursive least squares algorithms for total harmonic distortion reduction using shunt active power filter control", *Energies*, vol. 12, no. 8, pp. 1545, 2019 (DOI: 10.3390/en12081545).

[30] S. Haykin, "Adaptive filter theory (4-th edition)", Prentice Hall, 2001 (ISBN: 9780130901262).

[31] C. Cortes and V. Vapnik, "Support-vector networks", *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995 (DOI: 10.1007/BF00994018).

[32] B. Schölkopf, A. J. Smola, and F. Bach, "Learning with Kernels: Support vector machines, regularization, optimization, and beyond", MIT Press, 2002 [Online]. Available: https://alex.smola.org/papers/2001/SchHerSmo01.pdf

[33] B. Schölkopf, H. Ralf, and A. J. Smola, "A generalized representer theorem", *Int. Conf. on Comput. Learning Theory*, vol. 2111, Berlin, Springer, Heidelberg, pp. 416–426, 2001 (DOI: 10.1007/3-540-44581-1_27).

[34] K. Slavakis, B. Pantelis, and S. Theodoridis, "Online learning in reproducing kernel Hilbert spaces", *Signal Process. Theory and Machine Learning*, pp. 883–987, 2013 (DOI: 10.1016/B978-0-12-396502-8.00017-6).

[35] D. P. Wemerson, C. M. B. Jose, R. Cédric, and T. Jean-Yves, "Stochastic behavior analysis of the Gaussian kernel least-mean-square algorithm", *IEEE Trans. on Signal Process.*, vol. 60, no. 5, pp. 2208–2222, 2012 (DOI: 10.1109/TSP.2012.2186132).

[36] M. Kallas, "Méthodes à noyaux en reconnaissance de formes, prédiction et classification. Applications aux biosignaux", Ph.D. dissertation. Université de Technologie de Troyes, 2012 [Online]. Available: https://hal.archives-ouvertes.fr/tel-01088936 [in French].

[37] W. Liu, P. P. Pokharel, and J. C. Principe, "The kernel least-mean-square algorithm", *IEEE Trans. on Signal Process.*, vol. 56, no. 2, pp. 543–554, 2008 (DOI: 10.1109/TSP.2007.907881).

[38] P. Kumari and R. Wadhvani, "Wind Power Prediction Using KLMS Algorithm", in *Int. Conf. on Inventive Res. in Comput. App. (ICIRCA)*, Coimbatore, India, pp. 154–161, 2018 (DOI: 10.1109/ICIRCA.2018.8597419).

[39] K. Li and C. P. Jose, "No-trick (treat) kernel adaptive filtering using deterministic features", arXiv preprint arXiv:1912.04530, 2019 [Online]. Available: https://arxiv.org/pdf/1912.04530

[40] J. Ahmed, and K. A. Hamdi, "Spectral efficiency of asynchronous MC-CDMA with frequency offset over correlated fading", *IEEE Trans. On Vehicular Technol.*, vol. 62, no. 7, pp. 3423–3429, 2013 (DOI: 10.1109/TVT.2013.2253339).

[41] M. Zidane and R. Dinis, "A new combination of adaptive channel estimation methods and TORC equalizer in MC-CDMA systems", *Int. J. of Commun. Systems*, vol. 33, no. 11, pp. e4429, 2020 (DOI: 10.1002/dac.4429).

[42] M. S. Sakib and M. R. Ferdous, "Performance evaluation of MC-CDMA system for the next generation wireless networks", in *Int. Conf. on Innovation in Engin. and Technol. (ICIET)*, Dhaka, Bangladesh, pp. 1–5, 2018 (DOI: 10.1109/CIET.2018.8660783).

[43] S. Safi, M. Frikel, A. Zeroual, and M. M'Saad, "Higher order cumulants for identification and equalization of multicarrier spreading spectrum systems", *J. of Telecommun. and Informat. Technol.*, vol. 2, no. 1, pp. 74–84, 2011 [Online]. Available: https://www.itl.waw.pl/czasopisma/JTIT/2011/2/74.pdf

[44] V. Le Nir, J. M. Auffray, M. Hélard, J. F. Hélard, and R. Le Gouable, "Combination of Space-Time Block Coding with MC-CDMA Technique for MIMO systems with two, three and four transmit antennas", *IST Mobile & Wireless Commun. Summit*, Aveiro, Portugal, 2003 [Online]. Available: https://hal.archives-ouvertes.fr/hal-00005810/document

[45] S. Nobilet, "Etude et optimisation des techniques MC-CDMA pour les futures générations de systemes de communications hertziennes". Doctoral dissertation, INSA de Rennes, 2003 [Online]. Available: https://tel.archives-ouvertes.fr/tel-00004081/document [in French].



Rachid Fateh received his B.Sc. in Mathematics, Computer Science and Applications from Ibn Zohr University, Ouarzazate, Morocco in 2017 and M.Sc. in Telecommunications Systems and Computer Networks from Sultan Moulay Slimane University, Beni Mellal, Morocco in 2019. He is pursuing his Ph.D. in Mathematics and Computer Science from Sultan Moulay Slimane University, Beni Mellal, Morocco. His research interests include system identification, equalization of channel communications with a reproducing kernel Hilbert space, as well as signal processing and estimation.

https://orcid.org/0000-0002-0574-2105
E-mail: fateh.smi@gmail.com
Laboratory of Innovation in Mathematics, Applications and Information Technologies (LIMATI)
Multidisciplinary Faculty
Sultan Moulay Slimane University
Beni Mellal, Morocco



Anouar Darif received his B.Sc. in IEEA (Informatique Electrothecnique, Electronique and Automatique) from Dhar El Mahraz Faculty of Sciences at Mohamed Ben Abdellah University Fez, Morocco in 2005. He received the Diplôme d'Etudes Supérieurs Approfondies in Computer Sciences and Telecommunications from Faculty of Sciences in Rabat in 2007. He received his Ph.D. degree in Computer Sciences and Telecommunications from the Faculty of Sciences of Rabat in 2015. He is currently a Research and a Teaching Associate at the Multidisciplinary Faculty at University of Sultan

Moulay Slimane Beni Mellal, Morocco. His research interests include wireless sensor networks (WSN), mobile edge computing (MFC), Internet of Things (IoT) and Telecommunications (channel identification, IR-UWB, WCDMA, LTE, LTE-A,...).
https://orcid.org/0000-0001-8026-9189
E-mail: anouar.darif@gmail.com
Laboratory of Innovation in Mathematics, Applications and Information Technologies (LIMATI)
Multidisciplinary Faculty
Sultan Moulay Slimane University
Beni Mellal, Morocco



Said Safi received his B.Sc. degree in Physics (option: Electronics) from Cadi Ayyad University, Marrakech, Morocco in 1995, M.Sc. and Ph.D. from Chouaib Doukkali University and Cadi Ayyad University, in 1997 and 2002, respectively. He worked as a Professor of information theory and telecommunication systems at the National School for Applied Sciences, Tangier, Morocco, from 2003 to 2005. Since 2006, he has been a Professor of applied mathematics and programming at the Multidisciplinary Faculty, Sultan Moulay Slimane University, Beni Mellal, Morocco. In 2008, he received his Ph.D. degree in Telecommunication and Informatics from the Cadi Ayyad University. In 2015, he received the degree of Professor of Sciences at Sultan Moulay Slimane University. His general interests span the areas of communications and signal processing, estimation, time-series analysis, as well as system identification. Safi has published more than 50 journal papers and more than 70 conference papers concerned with the aforementioned topics. His research focuses currently on transmitter and receiver diversity techniques for single- and multi-user fading communication channels, as well as on broad band wireless communication systems.
https://orcid.org/0000-0003-3390-9037
E-mail: safi.said@gmail.com
Laboratory of Innovation in Mathematics, Applications and Information Technologies (LIMATI)
Multidisciplinary Faculty
Sultan Moulay Slimane University
Beni Mellal, Morocco

Paper

Phonetic Segmentation using a Wavelet-based Speech Cepstral Features and Sparse Representation Classifier

Ihsan Al-Hassani, Oumayma Al-Dakkak, and Abdlnaser Assami

Higher Institute for Applied Science and Technology HIAST, Damascus, Syria

https://doi.org/10.26636/jtit.2021.153321

Abstract—Speech segmentation is the process of dividing speech signal into distinct acoustic blocks that could be words, syllables or phonemes. Phonetic segmentation is about finding the exact boundaries for the different phonemes that composes a specific speech signal. This problem is crucial for many applications, i.e. automatic speech recognition (ASR). In this paper we propose a new model-based text independent phonetic segmentation method based on wavelet packet speech parametrization features and using the sparse representation classifier (SRC). Experiments were performed on two datasets, the first is an English one derived from TIMIT corpus, while the second is an Arabic one derived from the Arabic speech corpus. Results showed that the proposed wavelet packet decomposition features outperform the MFCC features in speech segmentation task, in terms of both F1-score and R-measure on both datasets. Results also indicate that the SRC gives higher hit rate than the famous k-Nearest Neighbors (k-NN) classifier on TIMIT dataset.

Keywords—Arabic speech corpus, ASR, F1-score, phonetic segmentation, sparse representation classifier, TTS, wavelet packet.

1. Introduction

The phonetic segmentation technique aims for identifying the starting and ending boundaries of each phoneme segment in continuous speech. This segmentation is crucial for creating phoneme databases used in text-to-speech (TTS) systems [1]–[3], and for transcribing speech corpus used in training hidden Markov models (HMMs) in ASR systems. Phonetic segmentation is also used in building a query-by-example (QbyE) spoken term detection (STD) application which is relatively a new application drawing increasing attention in recent years [4]. Knowledge of phoneme boundaries is also necessary in some cases of health-related research on human speech processing [4], such as diagnostic marker for Childhood Apraxia of Speech (CAS) [5]. Phonetic segmentation and annotation can be done either automatically or manually by expert phoneticians [6]. The main difficulty of this task is its subjectivity, because of the lack of distinct physiological or acoustic events that signal a phoneme boundary in some cases. In continuous speech, phoneme boundaries are sometimes difficult to locate due to glottalization, extremely reduced vowels or gradual decrease in energy before a pause [5]. As a result, there is no "correct" answer to the phoneme segmentation problem. Instead, a measure of the agreement between two alignments takes place, such as the agreement between two humans, or the agreement between human and machine [5]. Though manual segmentation is the most adequate [7] way for phonetic transcription. It suffers from being very tedious and time consuming, especially in the case of large speech corpora and spontaneous speech. In addition, manual segmentation suffers from labeler subjectivity and may not be able to maintain labeling consistency [8]. These difficulties stimulate the development of automatic phonetic segmentation techniques for continuous speech waveforms. These segmentation techniques are divided into two major categories: text-dependent (TD) and text-independent (TI) [9], [10]. In TD techniques, the phonetic annotation of the speech signal is already known and we need only to find the boundaries of each phoneme segment. Most text dependent segmentation techniques (also called explicit) are based on HMM with forced alignment Viterbi algorithm [9], [11]. On the other hand, TI segmentation methods (also called implicit or unsupervised) do not need any phonetic annotation for the speech signal to be segmented. Instead, they are generally based on sets of rules derived from encoding human knowledge to segment speech [12], like acoustic rate of change or other spectral variation metrics [13]–[15]. Such methods are called blind or model-free because they do not use modeling stage. Recently several studies proposed using different supervised and unsupervised machine learning techniques like ANN for phoneme segmentation [16], [17].

Sparse representation classifier [18]–[20] is relatively new machine learning technique that has demonstrated excellent performance in face recognition applications [18] and other applications [19], [20]. This classifier is based on extracting sparse code as discriminative features. Sainath *et al.* used Sparse coding for phoneme classification [21], [22] from test samples on a dictionary composed of phoneme exem-

plars as discriminative features, and fed these sparse codes to the sparse representation classifier. Sivaram *et al.* in [23] proposed employing sparse coding for phoneme recognition. They used the sparse code as a new speech feature to train multi-layer perceptron (MLP) network to get the posterior probabilities that will be used as emission likelihood of the HMM states. Every phoneme is modeled as a 3 state HMM and Viterbi decoder is used for phoneme recognition.

In this paper, we propose to use SRC for phoneme border detection and a speech parametrization algorithm based on the equivalent rectangular bandwidth (ERB) [24], [25] like wavelet packet decomposition entitled WP-ERB. The performance of the proposed classifier is compared to the k-NN classifier, and the performance of the proposed WP-ERB features are compared to the MFCC features in phoneme segmentation. In Section 2 we present related works for phoneme segmentation. The proposed phonetic segmentation system is described in Section 3. In Section 4 we present the conducted experiments and results. In Section 5 we summarize and conclude the paper.

2. Related Works

Significant work has been done on the problem of text independent speech segmentation. Some works used a set of rules derived from encoding human knowledge to segment speech [12], like acoustic rate of change, or other spectral variation metrics [13], [14]. Such methods are called model-free phonetic segmentation methods (also called metric-based or blind methods) because they do not incorporate any modeling strategy. Instead they rely on distance measures of the spectral changes among consecutive speech frames. These methods use the signal characteristics extracted in a signal analysis stage and a collection of thresholds to segment the signal [26]. The main issue with this approach is the difficulty to determine the optimal thresholds.

Javed *et al.* [27] proposed a strategy driven by cosine distance similarity scores for identifying phoneme boundaries. The proposed strategy helped in the selection of appropriate feature extraction technique for speech segmentation applications. Dusan in [28] investigated the use of spectral transition in segmentation, as he found high correlation between the maximum of the spectral transition and phoneme boundaries. The proposed method detects phoneme boundaries by looking for peaks in a spectral transition metric. Results showed an accuracy of 84.6% for frames of 20 ms TIMIT dataset, while no other performance metric was reported. Ramteke *et al.* [29] noted that in a well-spoken word, phonemes can be characterized by the changes observed in speech waveform. To get phoneme boundaries, Ramteke studied the signal level properties of speech waveform i.e. changes in the waveform during transformation from one phoneme to another. He addressed the problem of phoneme level segmentation from two aspects: segmentation of phonemes between voiced and unvoiced portions, and segmentation of phonemes within voiced and unvoiced regions. He used pitch and zero-frequency filter to get the region of change from voiced to unvoiced and vice versa. The segmentation of phoneme boundaries within voiced and unvoiced regions are approximated using the properties of power spectrum of correlation of adjacent frames of the signal. Finally, he proposed a finite set of rules on the variations observed in the power spectrum during phoneme transitions. The segmentation results of both approaches are combined to get the final phoneme boundaries. Three databases were used to test the proposed approach. An accuracy of 95.40%, 96.87% and 96.12% is achieved within the tolerance range of 10 ms respectively.

Recently several studies proposed using different supervised and unsupervised machine learning techniques to build a discriminative model that can be used in the phoneme segmentation task [4], [16], [17]. These methods are called model-based methods. Modeling stage is performed using either supervised or unsupervised approaches. Recently self-supervised learning algorithm was used for phoneme segmentation task [17].

In literature, research studies proposed various types of modeling approaches, like generalized gamma distribution model [30], graphical models [31], microcanonical multiscale formalism (MMF) [10], and acoustic segment modeling (ASM) [4]. Supervised and unsupervised machine learning techniques like ANN [16], [17], and genetic algorithm (GA) [32] were also used to learn the discriminative acoustic models.

Inspired by the success of using neural networks in speech recognition, different studies [16], [17] considered applying them to phoneme segmentation task. Different types of ANN were investigated. Dinler *et al.* [16] and Wang [33] suggested using gated recurrent unit (GRU) recurrent neural networks, while Kreuk [34] and Franke [35] proposed using bidirectional long-short term memory (LSTM) network. Lu [36] investigated the use of segmental recurrent neural network (RNN) for feature extraction. Lee [37] proposed using the cross-entropy loss with connectionist temporal classification loss in deep speech architecture for phoneme segmentation in view of performing speech synthesis. Wang [33] observed through experiments on the TIMIT corpus that GRU forget gate activations in trained recurrent acoustic neural networks correlate very well with phoneme which makes them preferable architecture for the task of boundary detection task. The advantage of both GRU and LSTM over standard RNNs lies in their ability to incorporate long temporal context information, and thus they give higher performance [16]. The GRU ensures the control of the information flow, similar to the LSTM unit, but without a need to utilize a memory unit [16]. The GRU has a simpler structure compared to standard LSTM models, and its popularity is gradually increasing [16].

Kreuk *et al.* [17] proposed a self-supervised representation learning (SSL) model for phoneme boundary detection. They proposed learning a feature representation from the raw waveform to identify spectral changes that match

phoneme boundaries accurately. For this task, they designed a convolutional neural network (CNN) to distinguish between pairs of adjacent frames and pairs of random distractor pairs. At test time, a peak detection algorithm is applied over the model outputs to produce the final boundaries [17]. Results show that the proposed SSL technique surpasses other unsupervised segmentation techniques.

All previous works used MFCC as acoustic features, though wavelet-based features has been shown to outperform MFCC in phoneme recognition application [25]. Also, no study has considered the application of the sparsity model in speech segmentation though it has achieved good success in many applications like noise robust ASR application [39], speech enhancement [40], [41], and speaker verification and identification [42]. In this work we propose using wavelet packet based acoustic features, as well as we examine the usefulness of sparse representation classifier SRC in phonetic segmentation task.

3. Proposed Method

The proposed segmentation system contains four stages: signal pre-processing, features extraction, dictionary creation, and phoneme segmentation. At the pre-processing stage silence from speech segments are removed, and pre-emphasis filter is applied to compensate for lips effects. Speech segments are then divided into overlapped frames of length 16 ms with 4 ms overlapping. After speech framing, acoustic features are being extracted: mel frequency cepstral coefficients (MFCC), and the proposed wavelet packet-ERP features [24], [25]. The block diagram of the proposed system is depicted in Fig. 1.

3.1. Wavelet Based Feature Extraction

The proposed wavelet packet feature extraction is based on the equivalent rectangular bandwidth (ERB) like wavelet packet decomposition proposed by Sahu in [25]. The whole frequency band is decomposed into 24 sub-bands according to the wavelet packet tree shown in Fig. 2 [25]. Once the WP decomposition is performed, energy in each frequency band is calculated, and the log of weighted energy is applied resulting in 24 cepstral coefficients. Discrete cosine transform (DCT) is then applied to decorrelate the 24 coefficients of filter bank energies, and variance feature (VF) of the 24 coefficients is also calculated. Finally, a total of 25 features are obtained for each frame.

Figure 3 illustrates the block diagram of the proposed wavelet packet-based feature extraction WP-ERP algorithm. We examined different types of wavelet filters with different degrees, such as Daubechies, coiflets and symlets filters. Experiments showed that the Coif5 filter gives the best performance in terms of segmentation accuracy.

3.2. Exemplar Dictionary Creation

At the dictionary creation stage, the feature vectors of the training phoneme/borders samples are warped together to form one matrix – the exemplar dictionary. The frame is labeled as phoneme (not a border, class 1) if either it does not contain a border (a border is a transition between two phonemes according to the manual annotation), or if most of the frame belongs to one phoneme, i.e. the border is not at the very start or the very end of the frame. On the other hand, the frame is labeled as border (class 2) if it contains a border between two phonemes and if the frame contains good percentage of both phonemes. The interval

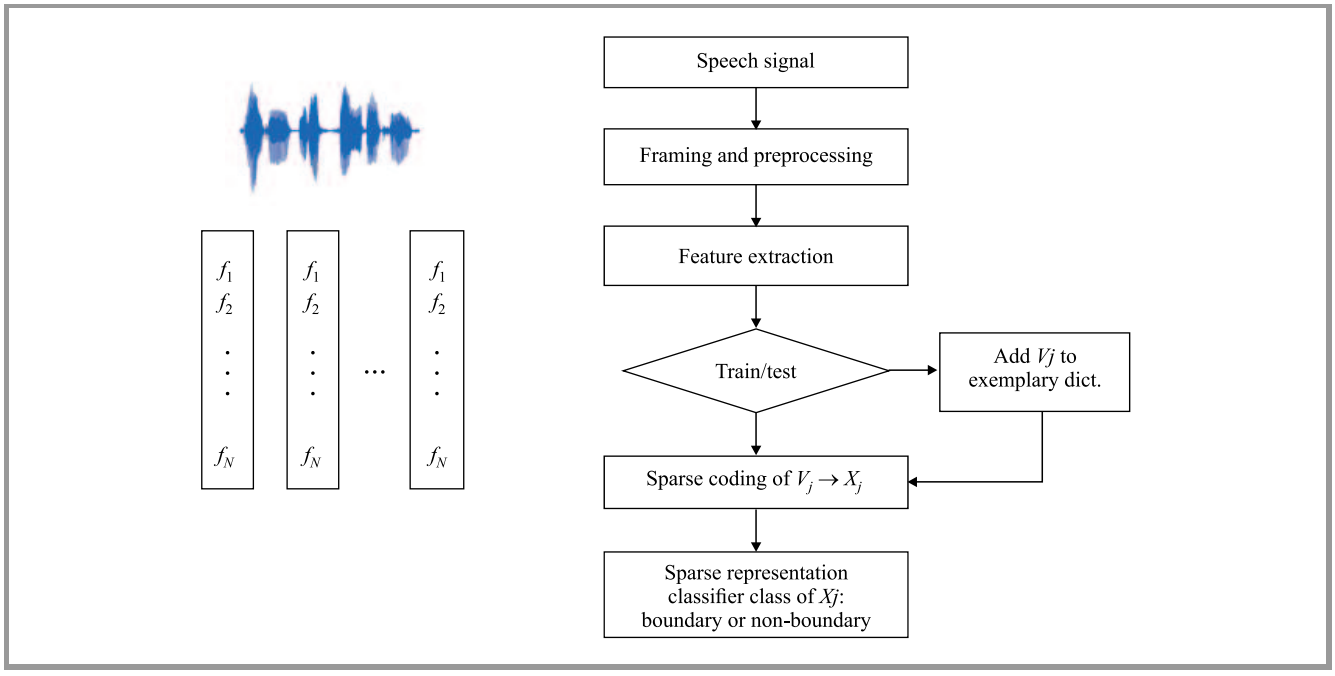


Fig. 1. Proposed phonetic segmentation system using SRC.

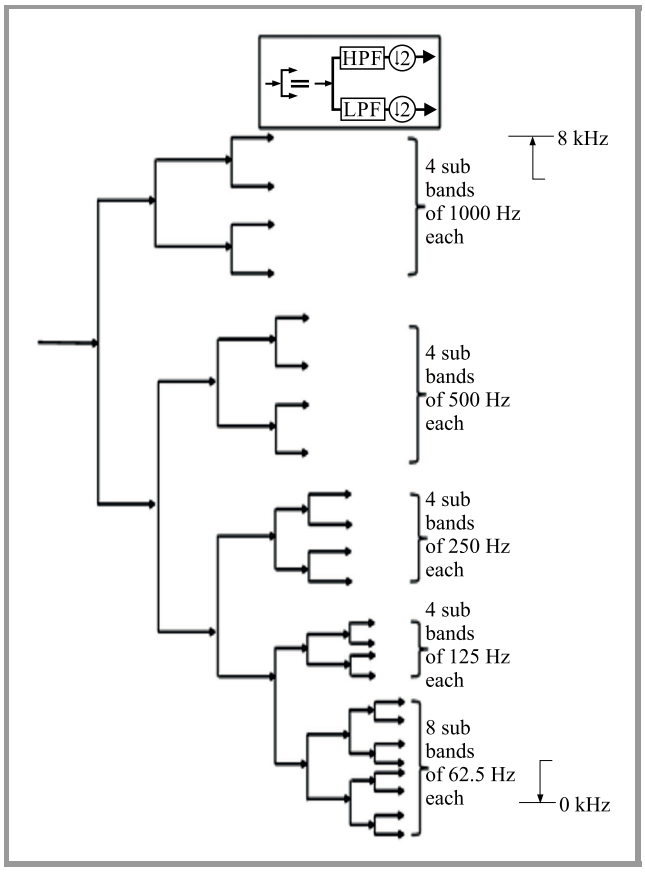


Fig. 2. 24 sub-band wavelet packet tree based on ERB scale.

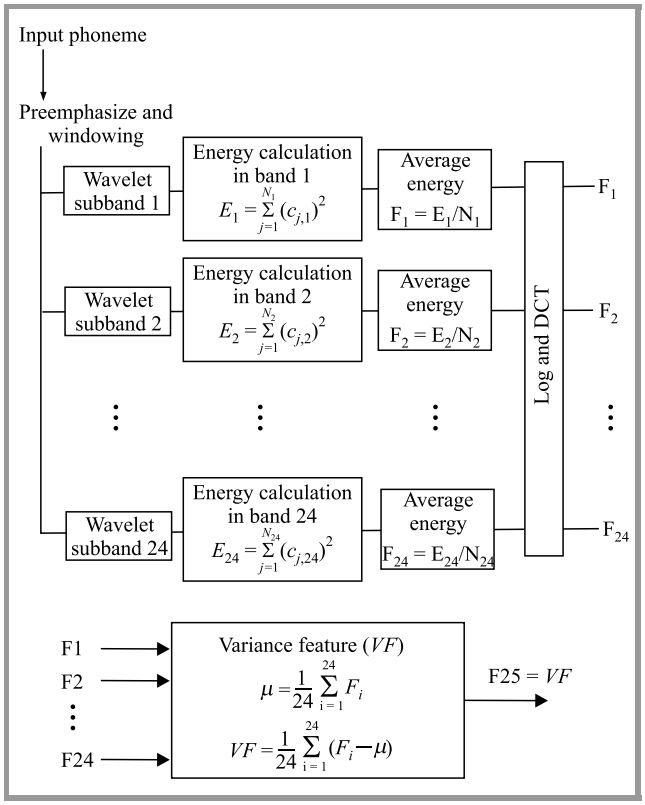


Fig. 3. Block diagram of the proposed wavelet-based WP-ERP feature extraction.

that defines the very start and the very end of the frame is taken equal to 3 ms, so that the tolerance interval for boundary detection is within $16 - 2 \times 3 = 10$ ms. Figures 4 and 5 depict the labeling strategy for both classes (border and not-border).

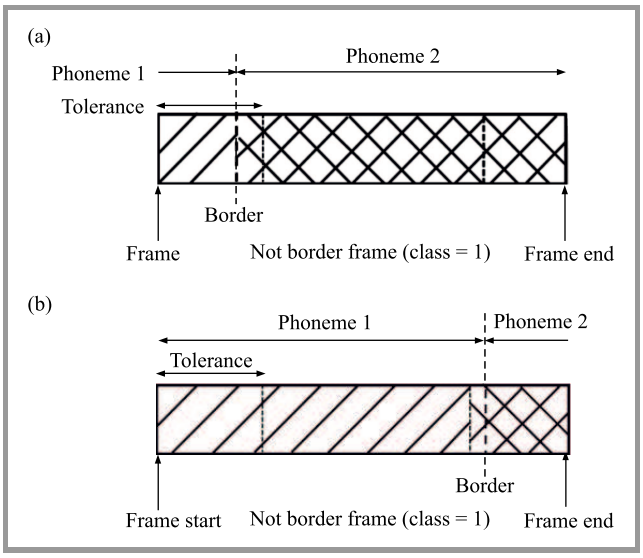


Fig. 4. Frame labeled as class 1 (not-border): the border is at the very start of the frame (a), the border is at the very end of the frame (b).

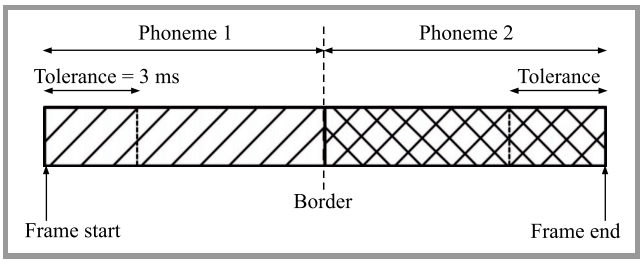


Fig. 5. Frame labeled as class 2 (border).

3.3. The Sparse Representation Classifier

The SRC is based on the sparse code, which can be defined as follows. Let $x \in \mathbb{R}^M$ be the signal that we want to encode sparsely. We suppose that there is a matrix (dictionary) $D \in \mathbb{R}^{M \times N}$, where $M \leq N$ such that y can be written as a linear combination of at most k columns of D, and $k \ll N$ is called the sparsity degree. The "sparse coding" problem [42] is to find the vector $a \in \mathbb{R}^N$ that contains only k non-zero elements such that:

$$(P0) \quad \min_{x \in \mathbb{R}^N} ||a||_\ell l_0 \quad \text{s.t.} \quad x = Da \ , \tag{1}$$

where $||a||_0$ is the l_0 pseudo-norm which represents the number of non-zero elements in a. The vector a which contains the sparse (few out of many) coefficients of the linear combination of the elements (called atoms) of the dictionary D, that represents the signal x is called the sparse

code. As D is over-complete (number of rows is less than the number of columns), this is an ill-posed inverse problem.

As the signal x might be corrupted with noise, the previous problem has been reformulated as:

$$(P_{0,\varepsilon}) \quad \min_{a\in\mathbb{R}^N} ||a||_0 \ \text{ s.t. } \ ||x-Da||_2 \leq \varepsilon \ . \tag{2}$$

This is a non-convex optimization problem that has been proved to be an NP-hard [38]. Replacing the l_0 norm by the l_1 norm "convexify" the problem and gives an equivalent problem called the basis pursuit denoising problem BPDN [39]:

$$(BPDN) \min_{a\in\mathbb{R}^N} ||a||_1 \ \text{ s.t. } \ ||x-Da||_2 \leq \varepsilon \ , \tag{3}$$

where $||a||_1 = \Sigma_1^N |a_i|$, and ε is some noise level energy. We can take the dual of the previous problem [42]:

$$(LASSO) \min_{a\in\mathbb{R}^N} ||x-Da||_2 \ \text{ s.t. } \ ||a||_1 \leq \tau \ , \tag{4}$$

where $\tau > 0$ is a regularization parameter, through which we control the sparsity degree (number of non-zero elements) of the sparse code a. This is called the "absolute shrinkage and selection operator" (LASSO) problem.

The SRC [18] works as follows. Having a training dataset that belongs to n class, for each class i the sub-dictionary D_i is formed by concatenating the corresponding training samples. These sub-dictionaries are wrapped together to form one dictionary D. The sparse code a of the feature vector for the test sample x is calculated using one of the sparse solvers available in the literature. For each class i, the selection operator $\delta_i(a)$ is applied on a, so that the elements of sparse code a corresponding to the sub-dictionary D_i are preserved, while all others are set to zero. Afterward, the linear approximation $D \times \delta_i(a)$ is calculated. SRC returns the class c that gives the closer approximation to the test sample x using the minimum Euclidean distance $||x - D \times \delta_i(a)||_2^2$.

Here, the SRC is calculated by:

1. Find the sparse code for the feature vector y, by solving Lasso – Eq. (4).
2. The class of x, is the index of sub-dictionary whose corresponding sub-sparse code energy is the highest:

$$c = \arg\min_i ||x - D \times \delta_i(a)||_2^2 \ ,$$

where $\delta_i(a)$ is a selector operator that selects the elements of sparse code a corresponding to the sub-dictionary D_i, and sets all others to zero.

In the literature [45] there are many algorithms that were developed for solving the previous sparse coding problems – Eqs. (2)–(3) and Eq. (4) – like: greedy orthogonal matching pursuit OMP, L1-minimization algorithms: GPRS, SPGL1, DALM, homotopy, L1LS. The Matlab implementations for these methods are available at [46], [47]. As many of the previous solvers include matrix inversion step, and due to the large size of the dictionary we used in our experiments, we could only use SPGL1 and OMP.

4. Results and Discussion

Experiments were conducted on two different datasets, the first one is an American English corpus derived from the TIMIT [49], and the second one is an Arabic one derived from the Arabic speech corpus [50].

TIMIT is one of the standards and phonetically balanced read speech English corpus, used in three domains: phoneme segmentation, phoneme classification and phoneme recognition systems to develop and evaluate the performance of these systems. This corpus consists of 6300 sentences recorded at 16 kHz rate with 16-bit sample, for the eight major dialects of American English spoken by 630 different speakers (438 males and 192 females), ten sentences for each [51]. These sentences are distributed in two sets, the training set with 4620 utterances from 462 speakers and the test set with 1344 sentences from 168 speakers. All sentences were segmented and labeled manually at the phoneme level.

TIMIT original transcriptions are based on 61 phonemes. Table 1 shows the TIMIT phoneme set, classified into voiced phonemes and unvoiced phonemes.

For experiments, a subset of 380 utterances from the complete set was used for training. Another subset of 100 utterances from complete test set was used for testing. We have excluded the "dialect" sentences (SA sentences) for both training and testing. Boundaries between two pauses, including stop closures, were also excluded from evaluation.

Arabic speech corpus [50] is a modern standard Arabic (MSA) speech corpus for speech synthesis. It contains phonetic and orthographic transcriptions of more than 3.7 hours of MSA speech aligned with recorded speech on the phoneme level. The annotations include word stress marks on the individual phonemes [52]. The corpus includes 1813 utterances recorded by a single speaker, with a 16-bit, 48 kHz speech waveform file for each utterance,

Table 1
TIMIT phoneme set (61 phonemes)

Voiced/ unvoiced	Type	Phonemes
Voiced	Vowels	iy, ih, eh, ey, ae, aa, aw, ay, ah, ao, oy, ow, uh, uw, ux, er, ax, ix, axr, ax-h
	Glides/ semi-vowels	l, r, w, y, hh, hv, el
	Stops	b, d, g
	Fricative	s, sh, f, th
	Nasal	m, n, ng, em, en, eng, nx
Unvoiced	Stops	p, t, k, dx, q
	Affricative	jh, ch
	Fricative	z, zh, v, dh
Pause and stop closures		pau, epi, h#, dcl, kcl, gcl, tcl, pcl, bcl

and a corresponding Praat [53] textgrid file for annotation. The annotation is based on a set of 82 Arabic phoneme. In out experiments, we used 200 utterances for training and another 100 utterances for testing. The wave files are down sampled to 16 kHz.

4.1. Performance Metrics

We used 6 common metrics (described below) that are generally used to assess phonetic segmentation algorithm. In the case of text independent segmentation (TI) techniques, the number of discovered segments might differ from the number of segments produced by manual segmentation. TI segmentation can be viewed as a boundary detection problem. The assessment of any detection algorithm is done by measuring: how much it truly detects, what should be detected, how much it truly rejects and what should be rejected:

Hit rate is a metric that measures how good an algorithm truly detects the goal. When a detected boundary matches corresponding boundary in the reference signal, this is called a hit rate (also called recall, RCL). It can be calculated as:

$$Hit\ rate = RCL = \frac{CDB}{ATB}\ 100\%\ , \tag{5}$$

where CDB is the number of correctly detected boundaries, and ATB is the number of all true boundaries. A hit rate of 100% implies that the algorithm is perfect in detecting boundaries, but it might detect non-boundaries frames and misclassify them as boundaries. For this issue, we use the precision measure.

Precision (PRC) is a metric that measures how precise the detection is, i.e. how good the algorithm is in detecting only what should be detected. It can be calculated as:

$$PRC = \frac{CDB}{CDB+IB} = \frac{CDB}{ADB}\ 100\%\ , \tag{6}$$

where ADB is the number of all detected boundaries (true and false) by the algorithm, and IB is the number of inserted boundaries (false detection). A precision of 100% means that the algorithm does not fire a false alarm which means detecting false boundaries. This is called over-segmentation error.

Specificity is a metric that measures how good the algorithm is in rejecting what should be rejected, and this is calculated as follows:

$$Specficity = \frac{AP-CDB}{AP-ATB}100\%\ , \tag{7}$$

where AP is the number of all points (frames in our case). We can see that a higher hit rate might come at the expense of lower specificity, and lower precision. Thus, hit rate and precision are not good metrics for assessing the overall performance of segmentation algorithm, as the increase in one of them might cause a decrease in the other. The overall objective effectiveness of the segmentation algorithm can be evaluated by three different measures: accuracy, the F1-score, and the R-measure [43].

Accuracy measures how accurate the algorithm is in both detection and rejection, and it is calculated by the formula:

$$Accuracy = \frac{CDB+CDNB}{AP}100\%\ , \tag{8}$$

where CDNB is the number of all true points detected as non-boundaries.

F1-score is the harmonic mean of recall and precision, which is used for assessing classification and prediction algorithms. It is calculated according to:

$$F1 = \frac{2\ PRC \times RCL}{PRC+RCL}\ . \tag{9}$$

F1-score takes its value in the unit interval between $0,\ldots,1$, where the score closer to 1 is better. A system with high recalls but low precision returns many results, but most of its predicted labels are incorrect. A system with high precision but low recall is just the opposite, returning very few results, but most of its predicted labels are correct. An ideal system with high precision and high recall will return many results, with all results labeled correctly [35].

Optimizing the operation of a speech segmentation algorithm is often a tradeoff between hit-rate and over-segmentation (or inversely, false-alarm rate and miss-rate) [48]. F1-score is one possible way to describe overall performance of an algorithm with a single value. However, F1-score is prone to stochastic hit-rate increases due to the over-segmentation issue [48].

R-value is a new distance measure proposed to describe performance using a single value that properly penalizes over-segmentation [48]. The optimal goal of segmentation is to achieve a hit-rate of 100% and an over-segmentation of 0%. This is called the target point (TP). The basis of the new metric is the algorithm's distance from TP and not the (hit-rate) gain achieved by over-segmentation.

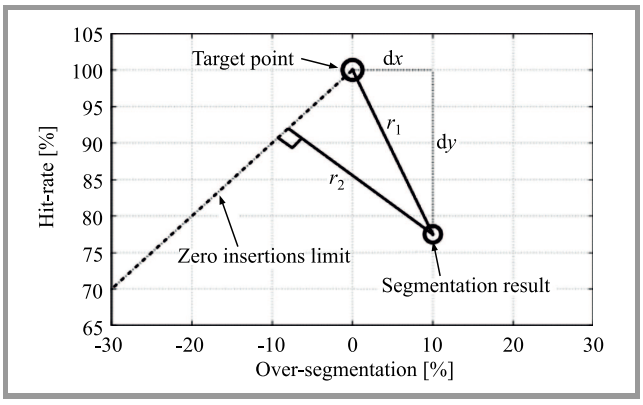


Fig. 6. Calculating R-measure [48].

On the segmentation performance plane illustrated in Fig. 6, a distance r_1 is derived (Eq. (10)) and a distance r_2 is measured (Eq. (11)), to appreciate the value of under-

segmentation compared to over-segmentation in the algorithm (i.e. less false positives).

$$r_1 = \sqrt{(100 - HR)^2 + OS^2}\ , \tag{10}$$

$$r_2 = \frac{1}{\sqrt{2}}\ (-OS + HR - 100)\ . \tag{11}$$

The distances r_1 and r_2 are then added together and normalized to have a maximum value of 1 at the target-point (Eq. (12)). This new distance measure is referred to as the R-value:

$$R = 1 - \frac{\text{abs}\,(r_1) + \text{abs}\,(r_2)}{200}\ . \tag{12}$$

R-value decreases as the distance to the target grows, similarly as F1-score does, but it makes more emphasis on over-segmentation by arguing that better hit rates might be achieved by simply adding random boundaries without any algorithmic improvement. This measure evaluates how close one is to the ideal segmentation R= 1.

4.2. Results

The performance of the proposed WP-ERB features was compared, against the performance of the well-known MFCC features. We have calculated the MFCC features using 24 mel coefficients[1], so that their dimensionality is close to that of the WP-ERB. All reported results are for boundary detection within a tolerance of 10 ms, and averaged over 10 random runs. Experiments were performed using a training set of 40,000 frames (20,000 frames for each class) and a test set of 6000 frames (3000 for each class). We tried different values for the sparsity regularization parameter τ in Eq. (4), and found that $\tau = 0.98$ gives that highest performance.

Table 2 shows the performance of the proposed segmentation system using the MFCC features and the proposed WP-ERB features, for both TIMIT and Arabic speech corpus. For TIMIT dataset, though MFCC achieves higher hit rate, but this at the expense of considerably lower specificity and lower precision. The overall performance of the proposed WP-ERB features gives the highest performance in terms of accuracy, F1-score, R-value, specificity and precision. The gain in terms of accuracy is about 2%, the gain in terms of precision and R-value is about 4%, and the gain in terms of specificity is about 12%, while the values of F1-score for both features are very close.

On the Arabic speech corpus, WP-ERB achieves a gain of 5% in terms of hit rate and 1.5% in terms of F1-score, over MFCC features, while the values of accuracy and R-value for both features are very close.

To assess the performance of SRC, we conducted the same segmentation scenarios using SRC, k-NN and SVM classifiers. We have not reported the results using SVM as it gave very bad performance, which can be explained by the fact that the two classes are mingled and cannot be separated by hyperplane. For k-NN tuning, we tried different values and found that $k = 80$ gives the best segmentation performance. Results are reported in Table 3.

On TIMIT dataset, we can see that the performance of the two classifiers are very close, which hints that the non-zero atoms that has the highest energies in the calculated sparse code are within the k-nearest points. SRC achieves higher hit rate at the expense of lower precision and specificity, but the overall performance in terms of F1-score is very close, though k-NN achieves a higher R-value and higher accuracy with a gain of about 2%. On the Arabic speech corpus k-NN achieves a better performance over SRC in terms of all performance parameters, with a mean gain of 1%.

Concerning the Arabic speech corpus, results show lower performance of about 2% than those on TIMIT in terms F1 score, and 6% lower in terms of hit rate. The modest results we obtained on the Arabic corpus are due to the consideration of 82 different phonemes. Some of these phonemes are so similar that they cannot be separated neither in time domain nor in the frequency domain, such as

[1]We tried different numbers for the mel coefficients of MFCC and all gave inferior performance.

Table 2
Segmentation performance: MFCC vs. WP-ERPC

Dataset	Feature	Accuracy	Hit rate	Specificity	Precision	F1-score	R-value
TIMIT	WP-ERB	**66.84**	81.60	**52.08**	**63.01**	71.11	**72.77**
	MFCC	64.60	**88.95**	40.25	59.82	**71.53**	68.86
Arabic speech corpus	WP-ERB	**65.17**	**75.55**	54.78	62.56	**68.45**	74.05
	MFCC	65.09	70.50	**59.67**	**63.62**	66.88	**74.15**

Table 3
Segmentation performance: SRC vs. k-NN classifier

Dataset	Classifier	Accuracy	Hit rate	Specificity	Precision	F1-score	R-value
TIMIT	SRC	64.67	**83.38**	45.97	60.68	70.24	70.63
	k-NN	**66.84**	81.60	**52.08**	**63.01**	**71.11**	**72.77**
Arabic	SRC	64.33	74.10	54.57	61.99	67.51	72.72
	k-NN	**65.17**	**75.55**	**54.78**	**62.56**	**68.45**	**74.05**

a phoneme and the geminated version of it. Also, in some vowels, as it is not possible to separate between the short version and the long version of the same phoneme (e.g. phoneme "a" and the phoneme "aa") if they are adjacent to each other.

We have examined different wavelet filter for extracting the proposed WP-ERB features. Results for TIMIT dataset are reported in Table 4, which shows that Coif5 gives the highest performance in terms of R-value and accuracy. The Sym8 filter gives the highest performance in terms of hit rate, with F1-score very close to this of Coif5, while a smaller R-value. Though the Haar wavelet achieves the highest specificity and precision but it has a low hit rate and thus low F1-score. In all, we can see the Coif5 has the best segmentation performance. Coiflets are the only wavelet basis that has vanishing moments of the scaling function φ [54], which is related to the "goodness" of the approximation of high-resolution scaling coefficients [54].

To assess the segmentation performance of the proposed algorithm depending on the type of phoneme boundary, we have calculated the hit rate for 3 different boundary types:

- V-V – boundaries that separate two voiced phonemes,
- U-V – boundaries that separate an unvoiced phoneme and voiced phoneme,
- U-U – boundaries that separate two unvoiced phonemes.

Results on TIMIT dataset and using SRC classifier are reported in Table 5. We can see that the boundary that separates two unvoiced phonemes are the hardest to detect achieving the lowest hit rate, while the boundary that separates two voiced phonemes are the easiest achieving the highest hit rate.

To study the effect of the size of the training set on the segmentation performance, we conducted three experiments using three training sets of different sizes: 10,000, 20,000, and 30,000 frames for each class, results are reported in Table 6. We can see that a training set of size 20,000 frames gives the best performance and increasing the size to 30,000 does not improve the performance. This can be explained by the fact that increasing the size of the training set might result in overfitting.

Table 5
Segmentation performance of WP features for different boundary types

Boundary type	Hit rate
V-V	84.34
U-V	81.69
U-U	67.73

To study the effect of the sparse coding solver we used 2 different sparse solvers: the simple greedy orthogonal matching pursuit OMP solver and SPGL-LASSO solver. Results are reported in Table 7 using the proposed WP-ERP features and a training set of 20,000 frames. We can see that though both solvers give very close hit rates, but SPGL-LASSO has a considerable increase over OMP in all other performance metrics.

Finally, in Table 8 we compared the performance of the proposed algorithm against two state of the art (SOTA) supervised phoneme segmentation on TIMIT dataset: Kreuk *et al.* [34] and Frank *et al.* [35]. Though results suggest that the proposed algorithm is inferior to the SOTA models over all metrics, but this is due to the classifier performance, as the two studies uses neural networks as classifier. The key result of this research is to show that the proposed wavelet based acoustic features outperform MFCC in the task of speech segmentation which was verified using the famous classifier k-NN and the proposed SRC.

Table 4
Segmentation performance of WP features using different wavelet filters on TIMIT dataset

Wavelet filter	Classifier	Accuracy	Hit rate	Specificity	Precision	F1-score	R-value
Sym8	SRC	64.99	**85.36**	44.62	60.65	70.91	70.27
	k-NN	66.45	82.90	50.01	62.38	**71.19**	72.08
Sym6	SRC	64.59	83.74	45.43	60.55	70.28	70.46
	k-NN	64.75	80.77	48.73	61.17	69.62	71.41
Haar	SRC	61.19	62.95	59.42	60.81	61.86	72.27
	k-NN	62.70	59.56	**65.83**	**63.55**	61.49	71.37
DB12	SRC	64.64	83.47	45.80	60.63	70.24	70.58
	k-NN	66.51	80.89	52.14	62.83	70.72	72.72
DB8	SRC	64.38	82.48	46.29	60.56	69.84	70.67
	k-NN	66.24	80.19	52.29	62.70	70.38	72.70
Coif5	SRC	64.67	83.38	45.97	60.68	70.24	70.63
	k-NN	**66.70**	81.24	52.15	**62.94**	70.93	**72.76**

Table 6
Segmentation performance using training sets for different size on TIMIT dataset

Size of training set	Features	Accuracy	Hit rate	Specificity	Precision	F1-score	R-value
10,000	WP-ERB	50.35	49.77	50.93	50.35	50.06	64.48
	MFCC	56.53	66.50	46.57	46.57	55.47	68.22
20,000	WP-ERB	**66.84**	81.60	**52.08**	**63.01**	71.11	**72.77**
	MFCC	64.60	**88.95**	40.25	59.82	**71.53**	68.86
30,000	WP-ERB	64.55	79.40	49.71	61.22	69.13	71.62
	MFCC	60.63	79.76	41.50	57.69	66.95	68.74

Table 7
Segmentation performance of WP features for different boundary types

Solver	Accuracy	Hit rate	Specificity	Precision	F1-score	R-value
OMP	59.46	77.61	41.30	56.93	65.68	68.42
SPGL-LASSO	**66.70**	**81.24**	**52.15**	**62.94**	**70.93**	**72.76**

Table 8
Comparison of phoneme segmentation models using TIMIT dataset

Model	Hit rate	Precision	F1-score	R-value	Tolerance
Kreuk *et al.* [34]	90.46	94.03	92.22	92.79	20 ms
Frank *et al.* [35]	88.10	91.10	89.60	90.80	20 ms
Proposed	66.70	62.94	70.93	72.76	10 ms

5. Conclusion

In this paper we proposed a new phonetic segmentation method based on speech parametrization technique entitled WP-ERB and sparse representation classifier. Results show that the proposed wavelet packet-based features outperform the classical MFCC features in speech segmentation task in terms of segmentation accuracy, precision, F1-score, and R-measure. The proposed WP-ERB features achieve a gain of about 4% in R-value and 2% in accuracy over MFCC on TIMIT dataset. On Arabic speech corpus the proposed WP-ERB features achieves a gain of 1.5% in terms of F1-score and 5% in terms of hit rate. We have also shown that using the SRC in phonetic segmentation achieves a higher hit rate over k-NN classifier on TIMIT dataset at the expense of lower precision and specificity, while no gain is achieved in terms of F1-score and R-value.

We think the moderate results with the Arabic corpus is due to the large number of considered phonemes (about twice the number of real phonemes). In later work, we will work on the Arabic dataset and merge the phonemes that cannot be separated and treat them as one phoneme (like geminated phonemes, short and long vowels of the same nature). As better results are obtained with TIMIT after phonemes merging, we expect the same for the Arabic corpus. This work is to be continued to see the effect of different dialects of the same language. TIMIT already contains many dialects, a comparative study will be undertaken to see the segmentation and classification performance on different dialects. On Arabic we try to collect data from other dialects, we expect some dialects far from Standard Arabic to be difficult to segment.

Speech style will also be an important point to study. As humans find sometimes difficulties in understanding some speech styles like fast speech or speech mixed with strong emotions. It will be interesting to see how far will differ the results with different speech styles.

The proposed phoneme segmentation system can further be improved by finding correlates between phonemes borders and prosodic features. Using those features together with acoustic knowledge of the phonemes, can be incorporated in a rule based to help increasing the system robustness.

References

[1] J. Glass, "A probabilistic framework for segment-based speech recognition", *Computer Speech & Language*, vol. 17, no. 2–3, pp. 137–152, 2003 (DOI: 10.1016/S0885-2308(03)00006-8).

[2] D. T. Chappell and J. Hansen, "A comparison of spectral smoothing methods for segment concatenation based speech synthesis", *Speech Commun.*, vol. 36, no. 3–4, pp. 343–373, 2002 (DOI: 10.1016/S0167-6393(01)00008-5).

[3] J. Adell and A. Bonafonte, "Towards phone segmentation for concatenative speech synthesis", in *Proc. of the 5th ISCA Speech Synthesis Workshop (SSW5)*, Pittsburgh, PA, USA, 2004, pp. 139–144 [Online]. Available: https://nlp.lsi.upc.edu/papers/adell04b.pdf

[4] H. Wang, T. Lee, C. Leung, B. Ma, and H. Li, "Acoustic Segment Modeling with Spectral Clustering Methods", in *IEEE/ACM Transac. on Audio, Speech, and Language Process.*, vol. 23, no. 2, pp. 264–277, 2015 (DOI: 10.1109/TASLP.2014.2387382).

[5] J. P. Hosom, "Speaker-independent phoneme alignment using transition-dependent states", vol. 51, no. 4, pp. 352–368, 2008 (DOI: 10.1016/j.specom.2008.11.003).

[6] J. P. van Hemert, "Automatic segmentation of speech", *IEEE Transac. on Signal Process.*, vol. 39, no. 4, pp. 1008–1012, 1991 (DOI: 10.1109/78.80941).

[7] A. Ljolje, J. Hirschberg, and J. P. H. van Santen, "Automatic speech segmentation for concatenative inventory selection", in *Progress in Speech Synthesis*, J. P. H. van Santen, R. W. Sproat, J. P. Olive, and J. Hirschberg, Eds., New York: Springer, 1997, pp. 304–311 (DOI: 10.1007/978-1-4612-1894-4_24).

[8] B. L. Pellom and J. H. L. Hansen, "Automatic segmentation of speech recorded in unknown noisy channel characteristics", *Speech Commun.*, vol. 25, no. 1–3, pp. 97–116, 1998 (DOI: 10.1016/S0167-6393(98)00031-4).

[9] G. A. Esposito, "Text independent methods for speech segmentation", *Nonlinear Speech Model. and App., Lecture Notes in Computer Sci.*, G. Chollet, A. Esposito, M. Faundez-Zanuy, M. Marinaro, Eds., Berlin, Heidelberg: Springer, 2005, vol. 3445 (DOI: 10.1007/11520153_12).

[10] V. Khanagha, K. Daoudi, O. Pont, and H. Yahia, "Phonetic segmentation of speech signal using local singularity analysis", *Digital Signal Processing*, vol. 35, no. C, pp. 86–94, 2014 (DOI: 10.1016/j.dsp.2014.08.002).

[11] D. T. Toledano, L. A. H. Gomez, and L. V. Grande, "Automatic phonetic segmentation", *IEEE Transac. on Speech and Audio Process.*, vol. 11, no. 6, pp. 617–625, 2003 (DOI: 10.1109/TSA.2003.813579).

[12] O. Scharenborg, V. Wan, and M. Ernestus, "Unsupervised speech segmentation: An analysis of the hypothesized phone boundaries", *J. of Acoustical Society of America*, vol. 127, no. 2, pp. 1084–1095, 2010 (DOI: 10.1121/1.3277194).

[13] B. D. Sarma and S. R. Mahadeva Prasanna, "Acoustic-Phonetic Analysis for Speech Recognition: A Review", *IETE Technical Review*, vol. 35, no. 3, pp. 305–327, 2017 (DOI: 10.1080/02564602.2017.1293570).

[14] M. Ziółko, J. Gałka, B. Ziółko, T. Drwięga, "Perceptual wavelet decomposition for speech segmentation", in *Proc. of the Interspeech, 11th Annual Conf. of the Int. Speech Commun. Association*, Makuhari, Chiba, Japan, 2010, pp. 2234–2237 (DOI: 10.21437/Interspeech.2010-614).

[15] D.-T. Hoang and H.-C. Wang, "Blind phone segmentation based on spectral change detection using legendre polynomial approximation", *The J. of the Acoustical Society of America*, vol. 137, no. 2, pp. 797–805, 2015 (DOI: 10.1121/1.4906147).

[16] Ö. Batur Dinler and N. Aydin, "An optimal feature parameter set based on gated recurrent unit recurrent neural networks for speech segment detection", *Appl. Sci.*, vol. 10, pp. 1273, 2020 (DOI: 10.3390/app10041273).

[17] F. Kreuk, J. Keshet, and Y. Adi, "Self-supervised contrastive learning for unsupervised phoneme segmentation", *Proc. of the Interspeech*, pp. 3700–3704, 2020 (DOI: 10.21437/Interspeech.2020-2398).

[18] J. Wright, A. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation", *IEEE Transac. on Pattern Anal. and Mach. Intell.*, vol. 31, no. 2, pp. 210–227, 2009 (DOI: 10.1109/TPAMI.2008.79).

[19] X. Gu, C. Zhang, and T. Ni, "A hierarchical discriminative sparse representation classifier for EEG signal detection", *IEEE/ACM Transac. on Comput. Biol. and Bioinform.*, vol. 18, no. 5, 2020 (DOI: 10.1109/TCBB.2020.3006699).

[20] Mh. Hajigholam, A. A. Raie, K. Faez, "Using sparse representation classifier (SRC) to calculate dynamic coefficients for multitask joint spatial pyramid matching", *Iranian J. of Sci. and Technol., Trans. Electr. Eng.*, vol. 45, pp. 295–307, 2021 (DOI: 10.1007/s40998-020-00351-3).

[21] T. N. Sainath, A. Carmi, D. Kanevsky, and D. Ramabhadran, "Bayesian compressive sensing for phonetic classification", *Proc. of the IEEE Int. Conf. on Acoustics Speech and Signal Process. (ICASSP)*, Dallas, TX, USA, 2010, pp. 4370–4373 (DOI: 10.1109/ICASSP.2010.5495638).

[22] T. N. Sainath and D. Kanevsky, "Sparse representations for speech recognition", A. Carmi, L. Mihaylova, S. Godsill Eds., book section "Compressed Sensing & Sparse Filtering", Berlin, Heidelberg: Springer, 2014, pp. 455–502 (DOI: 10.1007/978-3-642-38398-4_15).

[23] G. S. V. S. Sivaram, S. K. Nemala, M. Elhilali, T. D. Tran, and H. Hermansky, "Sparse coding for speech recognition", *IEEE Int. Conf. on Acoustics Speech and Signal Process. (ICASSP)*, Dallas, TX, USA, 2010, pp. 4346–4349 (DOI: 10.1109/ICASSP.2010.5495649).

[24] A. Bhowmick, M. Chandra, and A. Biswas, "Speech enhancement using Teager energy operated ERB-like perceptual wavelet packet decomposition", *Int. J. Speech Technol.*, vol. 20, pp. 813–827, 2017 (DOI: 10.1007/s10772-017-9448-7).

[25] P. K. Sahu, A. Biswas, A. Bhowmick, and M. Chandra, "Auditory ERB like admissible wavelet packet features for TIMIT phoneme recognition", *Engineer. Sci. and Technol., an Int. J. (Elsevier)*, vol. 17, no. 3, pp. 145–151, 2014 (DOI: 10.1016/j.jestch.2014.04.004).

[26] H. Frihia and Ha. Bahi, "HMM/SVM segmentation and labelling of Arabic speech for speech recognition applications", *Int. J. of Speech Technol.*, vol. 20, no. 3, pp. 563–573, 2017 (DOI: 10.1007/s10772-017-9427-z).

[27] M. Javed, M. M. A. Baig, and S. A. Qazi, "Unsupervised phonetic segmentation of classical Arabic speech using forward and inverse characteristics of the vocal tract", *Arab. J. Sci. Eng.*, vol. 45, pp. 1581–1597, 2020 (DOI: 10.1007/s13369-019-04065-5).

[28] S. Dusan and L. Rabiner, "On the relation between maximum spectral transition positions and phone boundaries", *Proc. of INTERSPEECH/ICSLP*, Pittsburgh, PA, USA, 2006, pp. 645–648 [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.569.3209&rep=rep1&type=pdf

[29] P. B. Ramteke and S. G. Koolagudi, "Phoneme boundary detection from speech: a rule based approach", *Speech Commun.*, vol. 107, no. 3, pp. 1–17, 2019 (DOI: 10.1016/j.specom.2019.01.003).

[30] G. Almpanidis, "Phonemic segmentation using the generalised Gamma distribution and small sample Bayesian information criterion", *Speech Commun.*, vol. 50, no. 1, pp. 38–55, 2008 (DOI: 10.1016/j.specom.2007.06.005).

[31] P. Teng, X. Liu, and Y. Jia, "Text-independent phoneme segmentation via learning critical acoustic change points", *Intell. Sci. and Big Data Engineer., Lecture Notes in Computer Sci.*, Berlin, Heidelberg: Springer, 2013, vol. 8261 (DOI: 10.1007/978-3-642-42057-3_8).

[32] A. H. Abo Absa, M. Deriche, M. Elshafei-Ahmed, Y. M. Elhadj, and B. Juang, "A hybrid unsupervised segmentation algorithm for Arabic speech using feature fusion and a genetic algorithm", *IEEE Access*, vol. 6, pp. 43157–43169, 2018 (DOI: 10.1109/ACCESS.2018.2859631).

[33] Y.-H. Wang, Ch.-T. Chung, and H.-Y. Lee, "Gate activation signal analysis for gated recurrent neural networks and its correlation with phoneme boundaries", *Interspeech*, Stockholm, Sweden, 2017 (DOI: 10.21437/INTERSPEECH.2017-877).

[34] F. Kreuk, Y. Sheena, J. Keshet, and Y. Adi, "Phoneme boundary detection using learnable segmental features", *IEEE Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP 2020)*, Barcelona, Spain, 2020, pp. 8089–8093 (DOI: 10.1109/ICASSP40776.2020.9053053).

[35] J. Franke, M. Mueller, F. Hamlaoui, S. Stueker, and A. Waibel, "Phoneme boundary detection using deep bidirectional LSTMs", *Proc. of the Speech Commun.: 12. ITG Symp.*, Paderborn, Germany, 2016, pp. 1–5 (ISBN: 9783800742752).

[36] L. Lu, L. Kong, Ch. Dyer, N. A Smith, and S. Renals, "Segmental recurrent neural networks for end-to-end speech recognition", *Proc. of the Interspeech*, 2016, pp. 385–389 (DOI: 10.21437/Interspeech.2016-40).

[37] Y. H. Lee, J. Y. Yang, C. Cho, and H. Jung, "Phoneme segmentation using deep learning for speech synthesis", *Proc. of the Conf. on Res. in Adaptive and Convergent Systems*, Honolulu, HI, USA, 2018, pp. 59–61 (DOI: 10.1145/3264746.3264801).

[38] J. F. Gemmeke, T. Virtanen, and A. Hurmalainen, "Exemplar-based sparse representations for noise robust automatic speech recognition", *IEEE Transac. on Audio, Speech, and Language Process.*, vol. 19, no. 7, 2011, pp. 2067–2080 (DOI: 10.1109/TASL.2011.2112350).

[39] D. Baby, T. Virtanen, J. F. Gemmeke, and H. Van Hamme, "Coupled dictionaries for exemplar-based speech enhancement and automatic speech recognition", *IEEE/ACM Transac, on Audio, Speech, and Language Process.*, vol. 23, no. 11, pp. 1788–1799, 2015 (DOI: 10.1109/TASLP.2015.2450491).

[40] V.-H. Duong, M.-Q. Bui, and J.-Ch. Wang, "Dictionary learning-based speech enhancement, active learning – beyond the future", *IntechOpen*, 2019 (DOI: 10.5772/intechopen.85308)..

[41] M. Hasheminejad and H. Farsi, "Frame level sparse representation classification for speaker verification", *Multimedia Tools and App.*, vol. 76, pp. 21211–21224, 2017 (DOI: 10.1007/s11042-016-4071-1).

[42] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A Survey of sparse representation: algorithms and applications", *IEEE Access*, vol. 3, pp. 490–530, 2015 (DOI: 10.1109/ACCESS.2015.2430359).

[43] B. K. Natarajan, "Sparse approximate solutions to linear systems", *SIAM J. on Comput.*, vol. 24, no. 2, pp. 227–234, 1995 (DOI: 10.1137/S0097539792240406).

[44] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit", *SIAM J. on Scientific Comput.*, vol. 20, no. 1, pp. 33–61, 1999 (DOI: 10.1137/S1064827596304010).

[45] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems", *Proc. of the IEEE*, vol. 98, no. 6, 2010, pp. 948–958 (DOI: 10.1109/JPROC.2010.2044010).

[46] SPGL1: A solver for large-scale sparse reconstruction [Online]. Available: https://www.cs.ubc.ca/ mpf/spgl1/index.html

[47] Matlab Benchmark Scripts, L-1 Benchmark Package [Online]. Available: http://people.eecs.berkeley.edu/∼yang/ software/l1benchmark/l1benchmark.zip

[48] O. J. Räsänen, U. K. Laine, and T. Altosaar, "An improved speech segmentation quality measure: the R-value", *Interspeech, 10th Annual Conf. of the Int. Speech Commun. Association*, Brighton, United Kingdom, pp. 1851–1854, 2009.

[49] J. S. Garofolo *et al.*, "TIMIT acoustic-phonetic continuous speech corpus", 1993 (DOI: 10.35111/17gk-bn40).

[50] Arabic Speech Corpus: a single-speaker, Modern Standard Arabic speech corpus made for high quality speech synthesis [Online]. Available: http://www.arabicspeechcorpus.com

[51] C. Lopes and F. Perdigao, "Phoneme recognition on the TIMIT database, speech technologies", *IntechOpen*, 2011 (DOI: 10.5772/17600).

[52] N. Halabi, "Modern standard Arabic phonetics for speech synthesis", *University of Southampton, Electronics & Computer Sci.*, Ph.D. Thesis, 2016 [Online]. Available: https://eprints.soton.ac.uk/ 409695/1/Nawar_Halabi_PhD_Thesis_Revised.pdf

[53] Praat: doing phonetics by computer [Online]. Available: https://www.fon.hum.uva.nl/praat/

[54] C. S. Burrus and J. E. Odegard, "Coiflet systems and zero moments", *IEEE Transac. on Signal Process.*, vol. 46, no. 3, pp. 761, 1998 (DOI: 10.1109/78.661342).



Ihsan Al-Hassani is a Ph.D. student at Higher Institute for Applied Science and Technology HIAST, Damascus, Syria. She received her M.Sc. degree in Telecommunication Engineering in 2013 from HIAST. Her research interests are focused on digital signal processing based embedded systems specialized in speech processing and digital communication.
E-mail: ihsan.alhassani@hiast.edu.sy
Higher Institute for Applied Science and Technology HIAST
Damascus
Syria



Oumayma Al-Dakkak received her Ph.D. in Electronics System from Grenoble Institute of Technology (Institut Polytechnique de Grenoble) in 1988. She is a Research Director and Head of Telecommunication Department, Higher Institute for Applied Science and Technology HIAST, Damascus, Syria. Her research interests include signal processing, speech processing, speech recognition, machine learning, cryptography and digital communication.
https://orcid.org/0000-0002-8842-0979
E-mail: oumayma.dakkak@hiast.edu.sy
Higher Institute for Applied Science and Technology HIAST
Damascus
Syria



Abdlnaser Assami received the M.Sc. degree from the ENST, Paris, France in 1996 and the Ph.D. degree from the University of Cergy-Pontoise, France in 2009. He is currently an Assistant Professor and Head of Digital Electronics Lab at the Higher Institute for Applied Sciences and Technology, Damascus, Syria. His research interests include man-machine communication, mobile communication, and Turbo-processing and its application to communication systems.
https://orcid.org/0000-0002-2036-5264
E-mail: abdelnasser.assimi@hiast.edu.sy
Higher Institute for Applied Science and Technology HIAST
Damascus
Syria

Paper

Research Concerning Electromagnetic Emissions from Residential On-grid PV Systems

Krzysztof Maniak, Janusz Sobolewski, Tomasz Tomczyk, Karolina Spalt, and Jacek Wroński

National Institute of Telecommunications, Poland

https://doi.org/10.26636/jtit.2021.157021

Abstract—The paper presents the results of grid power quality tests and wide frequency electromagnetic radiation level tests to which home power networks with photovoltaic systems have been subjected. The said results are meaningful from the point of view of safety and compatibility of electrical devices and the power supply system they are connected to. The tests covered stability of phase voltage and its frequency, harmonic level flicker, as well as RF radiated and emission levels. The measurements performed provide an answer to the question concerning the level of compliance of randomly selected home-installed PV systems with applicable recommendations and regulations concerning electromagnetic compatibility. The ability to meet the applicable standards translates into the health and safety of building inhabitants, which is the ultimate goal. Legal regulations concerning electromagnetic compatibility of renewable energy sources are consistent throughout the entire European Union.

Keywords—EMC, home-installed micro PV systems, photovoltaic power conversion, power grid quality, radiated emissions.

1. Introduction

The dynamic growth of global industry results in a significant increase in demand for energy. As it may be transmitted and transformed easily, electricity plays a major role in the process. Most power plants continue to generate energy by burning coal and lignite. This activity is accompanied by emissions of dust and carbon dioxide CO_2 [1]. Global CO_2 emissions equals 37.5 Gt [2] annually and have fallen slightly in 2020 due to rising coal prices [3]. If the situation persists, the estimated decrease in emissions could reach 2% per year. This fact causes, among other things, a worsening greenhouse effect and increases the average temperature of the globe.

Many outcomes of climate change caused by the use of coal are irreversible [4]. Numerous initiatives are being taken to limit these adverse phenomena [5]. For example, the World Meteorological Organization (WMO) aims to limit global average climate warming to 1.5°C by reducing CO_2 emissions by 2030 [6]. Therefore, it is desirable from the environmental point of view to look for alternatives to fossil fuels, such as generated by wind power, water and solar radiation. Geothermal waters are used as an energy source as well [6]. Energy generated by water [7] has been known and used for a long time and accounts, globally, for 19% of electricity production.

In individual households, photovoltaic installations are becoming extremely popular. With the continued advances in the field of solar installations taken into consideration, it is predicted that the world's total energy production from these sources will exceed 1 TWh by the end of 2022 [8]. With the trends observed in the European Union in mind, one has to conclude that it is necessary to come up with even better legal regulations favoring green electricity sources [9]. The same applies to Poland as well [10]. In Poland, 80% of all sunny days are observed in during the spring and summer months, i.e. from April to August. The annual number of sunny hours is about 1,600 and the average power of solar energy per 1 m^2 of terrain ranges from 970 to 1070 W/m^2 [11]. Due to the high power of solar systems, there is a risk of undesirable electromagnetic issues to which household electrical appliances are subjected. The said radiation impacts the health of the residents as well. Interference caused by the inverter converting DC voltage coming from the photovoltaic panels into AC voltage of the power grid is the main risk factor here. There are also reports in the literature that panels of which the PV systems are made act as antennas releasing signals into their surroundings. Such aerials may emit undesired RF signals related to the operation of the inverter [12], [13]. PV panels may also act as antennas receiving signals from the environment [14], [15]. Therefore, it is worth checking the degree to which residential solar power systems meet the requirements of applicable electromagnetic compatibility regulations. Despite the fact that the individual components of a solar system possess certificates confirming their acceptable levels of radiated and conducted emissions [16], [17], the entire systems – due to the fact that they comprise numerous components that fails to comply with the commonly applied practices and recommendations – may act as sources of excessive EM emissions.

There are no reports in the literature on comprehensive EMC studies of residential solar systems. According to

the authors, such studies are capable of providing valuable input on the compliance of complex solar installations with the requirements imposed by applicable standards.

2. Types of PV Installations for Domestic Use

The most common renewable energy installations for home use rely on solar energy. They form a part of the building's power network and are connected to the public power grid (hence, they are referred to as on-grid systems). The capacity of the system in an important criterion impacting its connection to the power grid. In the European Union, the following types of installations are distinguished for all renewable energy sources [18]:

- micro-installations – with rated power below 50 kW, operating at low voltage levels of 230/400 VAC. The surplus energy is fed to the grid via the user's connection;
- small installations – with rated power between 50 and 500 kW, connected to grids whose voltage levels are lower than 110 kVAC;
- large installations – with a power rating of over 500 kW. These systems require a connection to the main power supply point.

Solar power plants with the capacity of less than 50 kW, i.e. micro-installations, are the most common solution used in detached houses. Strong market competition between companies specializing in the installation of such solar systems results in significant pricing-related pressure, thus leading to components of insufficient quality being used. Subsidies offered by national governments for installing solar cells are an important factor as well. With the above taken into consideration, solar systems stand the best chance of becoming a widespread solution among electricity generators relying on micro, small and large installations. An example of a residential photovoltaic system is shown in Fig. 1.

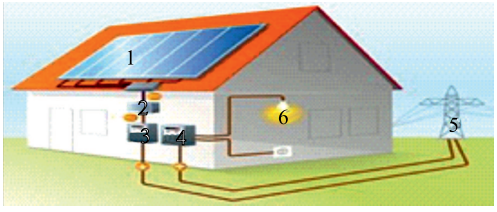


Fig. 1. Photovoltaic installation in a residential building: 1 – PV source of energy, 2 – inverter, 3 – feed-in, 4 – reference counter, 5 – power network, 6 – consumer appliances [19].

The DC voltage from the solar panels is converted into mains AC voltage by an inverter unit supplying energy to the consumer's appliances. The surplus electricity is fed back into the grid. Solar panels are the key component of any photovoltaic installation. Their peak power varies between 250÷400 W per unit. Three generations of such panels are available on the market [20], [21]. However, the first-generation panels are still most commonly used in modern photovoltaic installations due to their moderate costs and high efficiency of up to 22%. The lifetime of this type of cells is estimated at about 25 to 30 years.

3. Electromagnetic Compatibility of Photovoltaic Systems

With the constantly growing number of residential PV installations taken into account, it is necessary to examine their impact on the operation of other devices connected to the power network or operating in the vicinity of such a system. It is the inverter and its auxiliary equipment that is the main source of interference. The device is, in fact, a DC-AC converter capable of feeding energy into the power grid. The process of converting energy relies on switched topologies [22]. While the base frequency does not exceed 200 kHz, harmonic frequencies and, hence, the EM fields emitted reach, due to fast switching several tens of MHz. This process is responsible for electromagnetic emissions and also affects the parameters of the power network to which the inverter is connected. In addition, EM interference propagates through the power wiring that may act as parasitic antennas.

In the majority of micro installations, inverters with the output power rating of 1÷10 kW are used. Such a high electric power rating raises concerns about the potentially excessive levels of electromagnetic disturbance which may exert a negative impact on the safety of residents and on the proper operation of other electronic and electrical appliances. Despite the fact that all components of the installation meet the applicable safety-related and EM regulations, it is difficult to say whether the combined multi-module installation will ultimately continue to meet the requirements set forth in the relevant standards. This depends on numerous factors, i.e. on compliance with recommendations concerning the process of their manufacture, assembly methods, quality of electrical connections, compatibility of electrical parameters between specific installation components, etc. It is often the case that such recommendations are not followed by the technicians installing the system, which leads to its improper operation, to an increased level of disturbances and to the degradation of the parameters of the building's electricity network. This, in turn, may result in interference affecting Wi-Fi networks, in malfunctions of mobile telephony and in problems with radio and TV reception. The influence of EM fields on human health depends on field strength, frequency and exposure time [23]. With the applicable standards taken into consideration [16], [24], we shall attempt to evaluate specific PV systems installed in residential buildings, assessing such parameters of the grid as voltage, frequency, harmonic content, flicker, etc. Electromagnetic disturbance emissions will be evaluated as

well. For installations in residential buildings, the measurements shall also cover the strength of the 50 Hz EM field.

4. EMC Measurements

An accredited EMC testing laboratory dealing with in-situ measurements performed tests of solar systems installed in two single-family houses, both using first generation monocrystalline photovoltaic cells. Both buildings are new, with their electrical systems complying with the current standards [25]. Figure 2 shows the schematic of the completed systems for both buildings. In the first one (Object 1), 11 JAP60S10-275/SC panels were installed on the roof, each generating 275 W, meaning that the total peak power equals 3.0 kW. A single-phase SOFAR 2700TL-G3 inverter with a rated power of 3.7 kW and auxiliary instrumentation was located in a utility room adjacent to the garage. Data communication devices, such as Wi-Fi and GPRS modems were built into the inverter. At the other location (Object 2), 21 Longi LR4-60HPH-370M monocrystalline panels of 370 W each were mounted on the roof, providing a total power of 7.7 kW. A three-phase inverter of the Solar Edge SE8k type was installed in the garage of the building. The unit is additionally equipped with built-in data transmission modules, such as Wi-Fi, GPRS, and ZigBee. These generate additional emissions recorded during the measurements of radiated disturbances. Both homes were equipped in a similar manner, with standard mains-powered household appliances (washing machine, fridge, microwave oven, boiler etc.), television sets and computers, as well as LED interior lighting. In addition, internal Wi-Fi networks exist in both homes. During the measurements, all pieces of equipment which may cause additional electromagnetic disturbances, such as computers, TVs, chargers, i.e. devices containing AC/DC converters, were switched off. Importantly, in both cases the measurement probes were located at the main power connection point. The inverter and the junction box between the photovoltaic system and the household grid were also located there. Measurements were performed when the PV power generated was close to its maximum.

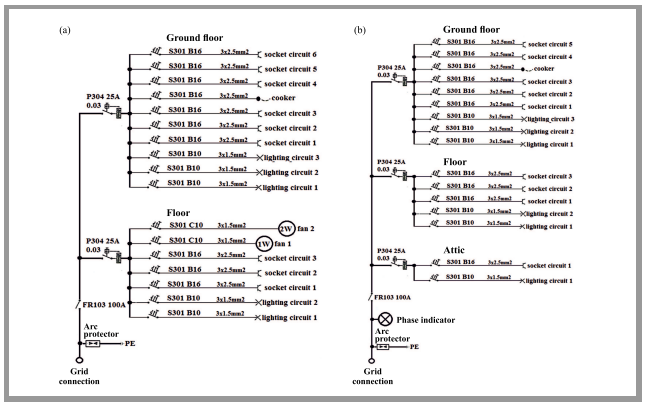


Fig. 2. Schematic diagram of the electrical system in Objects 1 (a) and 2 (b).

4.1. Power Network Quality Measurements

To assess the quality of the power network connected to a photovoltaic system, the criteria contained in standard [24] were applied. In both cases, power network parameters were measured using the Fluke 435 network quality analyzer. The measurement setup for the three-phase network is presented in Fig. 3, while Fig. 4 shows an image of the measuring station in Object 2.

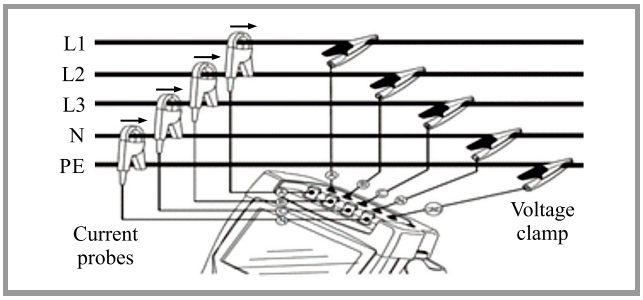


Fig. 3. Schematic diagram of the connection of the Fluke 435 power network analyzer to the grid.

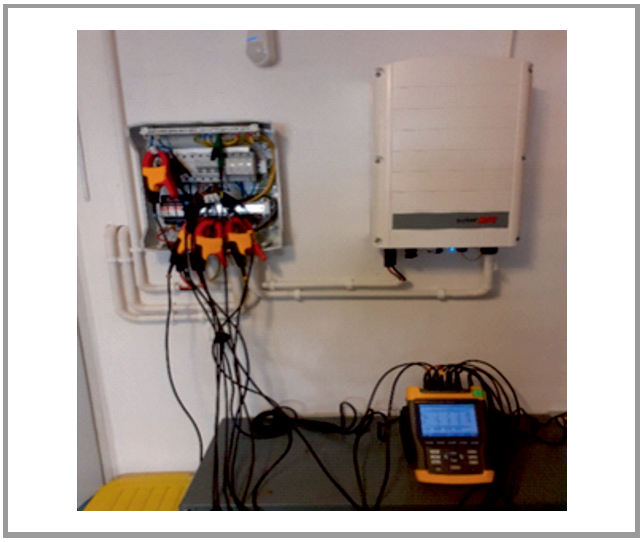


Fig. 4. Measurement station in Object 2.

The root mean square value of the network voltage is defined by [26]:

$$V_{rms} = \sqrt{\frac{1}{T_W}\Sigma_{n=0}^{T_W} {V_n}^2} \ , \tag{1}$$

where T_W is the measurement window with a duration of 10 periods of the network's AC voltage, V_n are the samples of the measured voltage.

The current measurement is carried out with the use of clamp probes. Such a solution ensures full safety due to the galvanic separation of the measurement circuit. For three-phase mains, the measurements were performed in a star

configuration. The rms value of the measured current is defined by [27]:

$$I_{rms} = \sqrt{\frac{1}{T_W}\Sigma_{n=0}^{T_W} {i_n}^2} \ , \tag{2}$$

where i_n are samples of the measured current.
The flicker index is measured by the P_{lt} parameter, i.e. is a long-term value measured over a period of 2 hours based on 12 measurements of 10-minute flicker. The P_{lt} value is determined from:

$$P_{lt} = \sqrt[3]{\Sigma_{i=1}^{12} \frac{{P_{sti}}^3}{12}} \ , \tag{3}$$

where P_{sti} is the 10-min flicker value.
According to standard [20], the measurement of total harmonic distortion in the mains voltage waveform should be a sum of first 40 harmonic components of the fundamental frequency [24]:

$$THD = \sqrt{\Sigma_{h=2}^{40} \left(V_h\right)^2} \ . \tag{4}$$

The voltage crest factor is defined by:

$$CF = \frac{V_{peak}}{V_{rms}} \ , \tag{5}$$

where U_{peak} is the amplitude of the voltage waveform. For a sinusoidal waveform, CF $= 1.414$.
The power network parameter values obtained at both Objects are summarized in Table 1. In addition, shifts between the power network phase voltages were recorded for Object 2 (Fig. 5).

Table 1
Basic power network parameters in both objects

Object 1	Object 2		
L1	L1	L2	L3
AC voltage [V]			
252.44	236.67	240.40	240.98
Voltage crest factor			
1.42	1.40	1.40	1.40
Mains frequency [Hz]			
49.992	49.996	-	-
THD [%]			
2.3	2.8	3.0	2.9
Flicker Plt			
0.31	0.24	0.37	0.83

Based on the measurements, one may observe that all relevant power network parameters are within the ranges specified in the standard [24]. The rms value of the AC voltage does not exceed the variation threshold of $\pm 10\%$, and frequency remains within the $\pm 1\%$ range. Similarly, THD of the mains voltage does not exceed the permitted 8% level and flicker does not exceed unity. Here, only the line voltage harmonics are included in measurements. This is

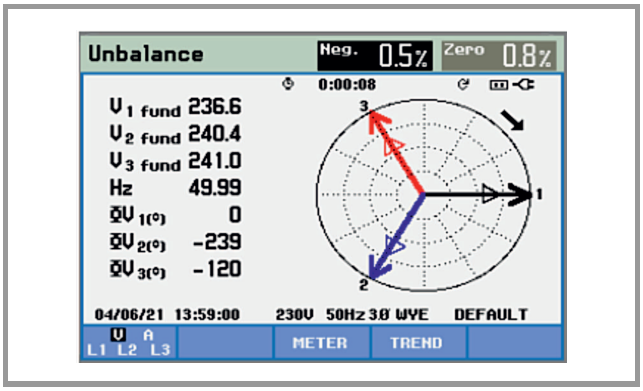


Fig. 5. Phase dependence of the 3-phase grid AC voltage at Object 2.

due to the fact that a sinusoidal voltage waveform is generated by the inverter circuit, as the PWM switching process significantly increases THD. On the other hand, the mains current's THD strongly depends on the grid load. For example, many switching power supply lines (without a PFC unit) have a diode bridge rectifier connected to a large capacitor at AC input. The current is then drawn in pulses, which causes significant distortion of waveforms and an increase in THD. Voltage relations and phase shifts for the three-phase grid in Object 2 are also in compliance with the related standard under which an asymmetry of up to 2% between phase voltages is permitted.

4.2. Radiated Emissions Measurements

The radiated emissions were examined within a frequency range that is crucial for the operation of electronic devices used at home and in accordance with standard [16]. The band under consideration is also important from the point of view of ensuring healthy living conditions for the residents. A biconical SAS-521F-7 antenna by A. H. Systems was used for V (vertical) and H (horizontal) polarity measurements. The tests were performed in the frequency range of 30 MHz÷6 GHz. In accordance with the recommendations set forth in standard [16], the antenna was placed

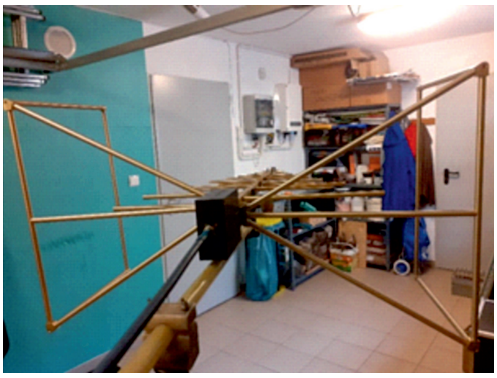


Fig. 6. Radiated emissions measurement setup – Object 2.

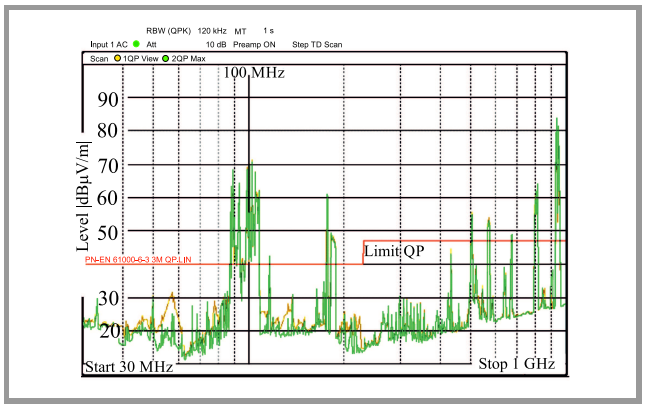


Fig. 7. Radiated emissions acquired at Object 1 for horizontal antenna polarity in the 30 MHz÷1 GHz band (green line – background level, yellow – QP emissions, red – applicable limits). (see the digital edition for color images)

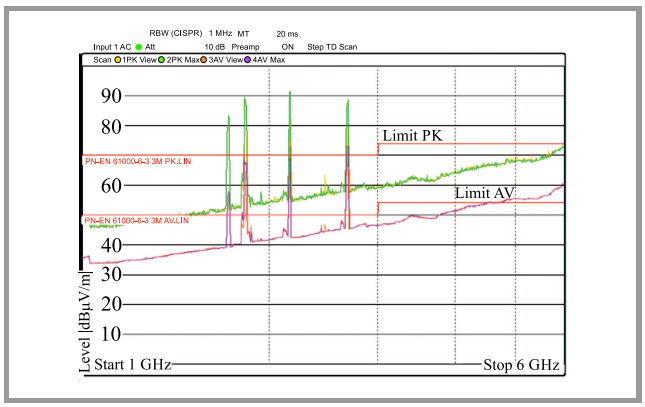


Fig. 8. Radiated emissions for Object 1, with horizontal antenna polarity in the 1 GHz÷6 GHz band (yellow – PK emissions, green line – background PK, orange – AV emissions, magenta – background AV, red – applicable limits).

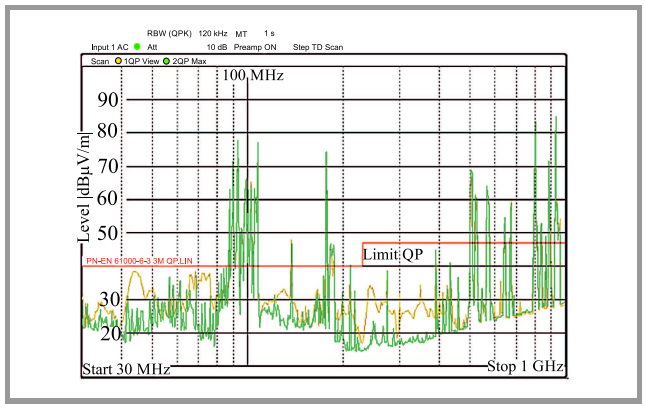


Fig. 9. Radiated emissions acquired at Object 2 for horizontal antenna polarity in the 30 MHz÷1 GHz band (green line – background level, yellow – QP emissions, red – applicable limits).

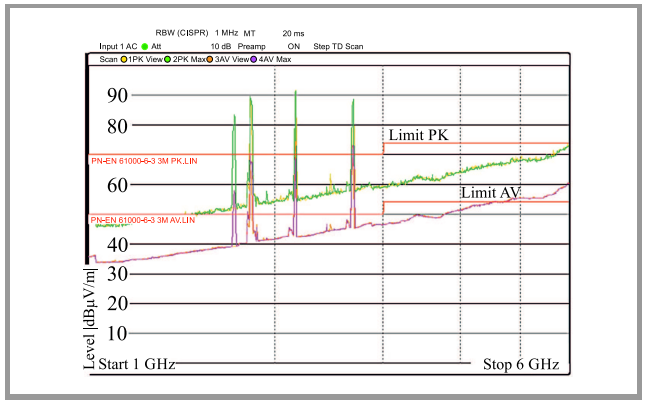


Fig. 10. Radiated emissions for Object 2, for horizontal antenna polarity in the 1 GHz÷6 GHz band (yellow line – PK emissions, green – background PK, orange – AV emissions, magenta – background AV, red – applicable limits).

at a distance of 3 m from the device undergoing the test (i.e. in the case of both locations, the inverter plus its auxiliary equipment). For the 30 MHz÷1 GHz band, standard [16] requires that a quasi peak (QP) detector be used. For the 1–6 GHz range, peak (PK) and average (AV) detectors were used. Measurements of the background level were performed with the inverter switched off, thus making it possible to observe EM signals present in the tested object and not related to the operation of the solar system. No significant variations were found in the results depending on antenna polarization. However, for both frequency ranges, results obtained with the measuring antenna in the horizontal configuration were considered. The measurements performed at both Objects are shown in Figs. 7–8 for Object 1, and in Figs. 9–10 for Object 2. The test bench used is shown in Fig. 6.

In the 30 MHz–1 GHz range we are interested with, areas could be identified at both Objects where the background signal considerably exceeded the permitted levels. For the 30 MHz–1 GHz band, main source of EM emissions includes FM radio waves which exceeded the permitted levels in the region around the 100 MHz mark. The second big source of EM emissions present in the background includes, in most cases, signals from terrestrial DVB-T television. This is visible near the 180 MHz mark and in the 400 MHz–1 GHz band. In higher portions of spectrum, i.e. 1÷6 GHz, one may notice signals of GSM transmitters: 1.8 GHz, 2.4 GHz, and 2.6 GHz. These may be eliminated from the measurement data. The radiated emissions generated by the inverter, i.e. the main source of EM disturbances, are within the limit allowed under the standard [16]. However, in Object 2 for instance, below the 80 MHz mark, the margin equals approx. 3–4 dB. This is visible for the waveform recorded in the vicinity of 45 and 60 MHz. The increase in radiated emissions is most probably caused by the operation of the inverter itself.

4.3. 50 Hz Electromagnetic Field Measurements

Due to the fact that the inverter is a source of 50 Hz voltage that (the same as in the power network) and works with high power, it is necessary to check whether the permissible levels of EM fields present in the proximity of this device

are not exceeded. The measurements were performed for residential building conditions, in accordance with standard [28]. For the 50 Hz mains frequency, the allowed values equal 5000 V/m for the electric field and 80 A/m for the magnetic component. Since no measurement distances from the tested device (inverter) are recommended, only the general recommendation was followed, namely the measurements were performed 0.3 m from the unit and 1.8 m above the floor. At this position, the maximum EM field strength was obtained for both Objects using a field intensity meter type HI-3604 by Holaday Industries. The location of the measurement sites is presented in Fig. 11, and a summary of the results obtained is given in Table 2.

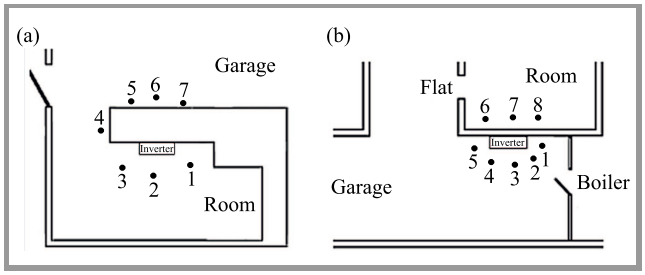


Fig. 11. Location of measurement points in Object 1 a) and Object 2 b).

During the measurements, power consumption of the household was close to maximum. The intensity of the relevant EM field components at each test point did not exceed the permitted level in close proximity to the inverter. Therefore, at other locations within the household, field strength values describing the mains frequency electromagnetic field will be correspondingly lower and thus safe for the environment.

Table 2
50 Hz electromagnetic field values for different measurement locations

No.	Object 1		Object 2		Evaluation
	Electric field [V/m]	Magnetic field [mA/m]	Electric field [V/m]	Magnetic field [mA/m]	
1	51.6	90.0	18.7	583.2	No issues
2	61.4	47.4	12.2	489.6	
3	44.2	39.3	47.3	432.0	
4	29.0	20.5	44.8	295.2	
5	36.5	17.3	9.8	316.8	
6	29.2	14.2	79.9	108.0	
7	35.5	22.2	52.8	236.2	
8	-	-	39.5	210.6	

5. Propagation of Radiated Disturbances

It should be noted that the strength of the electric field was measured in accordance with household-applicable standards, in close proximity to the source, i.e. the inverter and its auxiliary equipment. In such a scenario, the issue of spatial distributions of the emitted EM disturbances at longer distances from the source may arise. However, due to the complex propagation conditions prevailing in houses, it is difficult to determine unambiguous propagation paths of the disturbances emitted by the PV system's secondary components. The degree of susceptibility of electronic devices depends on their location in the household and it is difficult to precisely determine the position of the residents at a given time of the day. In this case, the propagation of the EM field in a homogeneous semi-conductive medium with variable characteristic parameters should be considered.

The source of EM disturbances originating from the inverter along with its auxiliary equipment, due to the fact that its size is small in relation to the entire building with the solar system installed, may be presented in the form of an electric dipole placed at the beginning of the Cartesian coordinate system and oriented along the z axis, as pictured in Fig. 12.

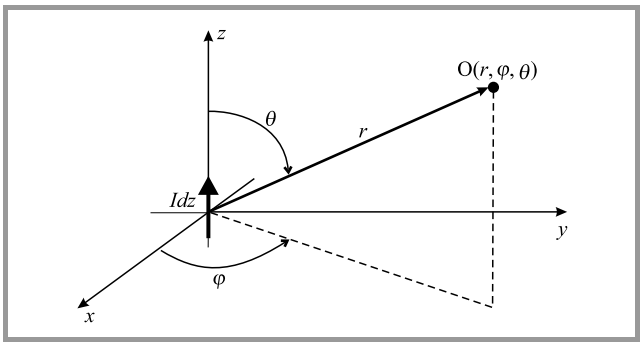


Fig. 12. Electromagnetic field of an electric dipole placed in a homogeneous semiconductive medium.

The semiconductive medium in which the electromagnetic field of the electric dipole is propagated is characterized by [29]:

- magnetic permeability $\mu = \mu_1 \cdot \mu_0$ [H/m],
- electric permeability $\varepsilon = \varepsilon_1 \cdot \varepsilon_0$ [F/m],
- electrical conductivity σ [S/m],
- specific electrical resistance $\rho = 1/\sigma$ [Ωm].

On this basis, it is possible to determine the important parameter of the medium, i.e. the wave number expressed by [29]:

$$\gamma = \sqrt{\frac{\mathrm{i}\omega\mu}{\rho} - \varepsilon\mu\omega^2} = \sqrt{\mathrm{i}\omega\mu\left(\sigma + \mathrm{i}\omega\varepsilon\right)}\,, \qquad (6)$$

where ω is the pulsation [rad/s].

The EM field components expressed in spherical coordinates for the observation point O (see Fig. 12) take the form [30]:

$$E_r = \frac{-Idz}{4\pi} \cdot 2\mathrm{i}\omega\mu \cos Q \frac{\mathrm{e}^{\mathrm{i}\gamma r}}{\gamma^2 r^3} (1 - \mathrm{i}\gamma r) \ , \tag{7}$$

$$E_\theta = \frac{-Idz}{4\pi} \cdot \mathrm{i}\omega\mu \sin Q \frac{\mathrm{e}^{\mathrm{i}\gamma r}}{\gamma^2 r^3} \left(1 + \gamma^2 r^3 - \mathrm{i}\gamma r\right) \ , \tag{8}$$

$$H_\varphi = \frac{Idz}{4\pi} \cdot \sin Q \frac{\mathrm{e}^{\mathrm{i}\gamma r}}{r^2} (1 - \mathrm{i}\gamma r) \ . \tag{9}$$

The values of field strength components depend not only on the specification of the propagation medium, but also on the frequency and distance from the source of RF energy. Under real conditions, it would be necessary to create a wideband computational model of each object or to analyze data using the relations given in Eq. (7)–(9). It may be assumed that the values of the EM field strength components decrease with distance. In this case, it will be necessary to perform bests of electric field strength in the vicinity of the expected source of emissions and to assess whether the obtained values do not exceed the levels permitted under the standard. At longer distances from the radiation source, field strength values will be lower than the ones that have been measured. This is because in an average interior, there are several surfaces reflecting and absorbing EM radiation [31], [32]. This causes a reduction in the levels of emitted fields, with the degree of such a reduction being a function of the distance from the source.

6. Conclusion

The paper attempts to assess whether domestic PV systems pose a risk to residents and to the electrical infrastructure, under normal conditions in which solar installations operate. In order to answer this question, extensive research related to electromagnetic compatibility needs to be conducted. In the case of electrical installations working together with renewable energy sources, tests of basic power network quality parameters are recommended, such as effective value of the mains voltage, frequency, THD, flicker and other factors, depending on the operating conditions prevailing at a given location. In addition, for the sake of the household members and to avoid malfunctioning of the electrical appliances, it is necessary to measure the RF emissions generated by the devices performing switched mode energy conversion. In photovoltaic systems, this includes the inverter and its auxiliary equipment. In order to ensure the safety of humans, it is important to measure the intensity level of 50 Hz mains fields. It is recommended to assess the acquired values based on standards applicable to residential installation. The test campaign performed by the authors and covering two randomly selected domestic PV micro-installations shows that the compliance of the power network with the recommendations set forth in the standards is ensured. Similarly, the measurement of high-frequency electromagnetic field emissions in the 30 MHz÷6 GHz range did not identify and situations in which the levels permitted under the standard would be exceeded. This means that the installations are safe from the point of view of their users. The measurements of specific electromagnetic field components at the mains frequency failed to identify any risks or additional threats that people could be subjected to as well.

Acknowledgement

The authors would like to thank Tomasz Gurdziałek for agreeing to perform the measurement campaigns of photovoltaic systems installed at his home.

References

[1] R. D. Bressler, "The mortality cost of carbon", *Nature Commun.*, no. 12, 2021 (DOI: 10.1038/s41467-021-24487).

[2] Emission Gap Report 2019, *United Nations Environment Programme* [Online]. Available: https://www.unep.org/resources/emissions-gap-report-2019

[3] J. F. Green, "Does carbon pricing reduce emissions? A review of ex-post analyses", *Environmental Research Letters*, vol. 16, no. 4, 2021 (DOI: 10.1088/1748-9326/abdae9).

[4] S. Solomona, G. K. Plattnerb, R. Knuttic, and P. Friedlingsteind, "Irreversible climate change due to carbon dioxide emissions", in *Proc. of the National Academy of Sci. of the United States of America*, vol. 106, no. 6, 2009 (DOI: 10.1073pnas.0812721106).

[5] F. L. Suocheng Dong, Z. Li, Y. L. Shantong Li, and Y. Wan, "The improvement of CO2 emission reduction policies based on system dynamics method in traditional industrial region with large CO2 emission", *Energy Policy*, vol. 51, pp. 683–695, 2012 (DOI: 10.1016/j.enpol.2012.09.014).

[6] V. Masson-Delmotte *et al.*, "Global warming of 1.5°C", IPCC Special Report, *Intergovernmental Panel on Climate Change*, 2019 [Online]. Available: https://www.ipcc.ch/sr15/download/

[7] V. Laxmi Kalyani, M. Kumari Dudy, and S. Pareek, "Green energy: the need of the world", *J. of Management Engineer. and Informat. Technol.*, vol. 2, no. 5, 2015 [Online]. Available: https://www.researchgate.net/publication/283482870_GREEN_ENERGY_The_NEED_of_the_WORLD

[8] S. Chowdhurya *et al.*, "An overview of solar photovoltaic panels' end-of-life material recycling", *Energy Strategy Reviews*, vol. 27, 2020 (DOI: 10.1016/j.esr.2019.100431).

[9] A. Pegels, G. Vidican-Auktor, W. Lu Tkenhorst, and T. Altenburg, "Politics of green energy policy", *J. of Environment & Development*, vol. 27, no. 1, pp. 26–45, 2018 (DOI: 10.1177/1070496517747660).

[10] J. Kiciński, "Green energy transformation in Poland", *Bull. of the Polish Academy of Sci. Technic. Sci.*, vol. 69, no. 1, 2021 (DOI: 10.24425/bpasts.2020.136213).

[11] A. Dobrzycki, D. Kurz, S. Mikulski, and G. Wodnicki, "Analysis of the impact of building integrated photovoltaics (BIPV) on reducing the demand for electricity and heat in buildings located in Poland", *Energies*, vol. 13, no. 10, 2020 (DOI: 10.3390/en13102549).

[12] H. Lgarashi and S. Suenaga, "Electromagnetic noise from solar cells", in *Proc. Conf. Record of the Thirty-first IEEE Photovoltaic Specialists Conf.*, Lake Buena Vista, FL, USA, 2005 (DOI: 10.1109/PVSC.2005.1488506).

[13] C. Jettanasen and C. Pothisarn, "Performance and electromagnetic interference mitigation of DC-DC converter connected to photovoltaic panel", *Int. J. of Smart Grid and Clean Energy*, vol. 8, no. 6, pp. 806–812, 2019 (DOI: 10.12720/sgce.8.6.806-812).

[14] C. Baccouch, D. Bouchouicha, H. Sakli, and T. Aguili, "Patch antenna based on a photovoltaic cell with a dual resonance frequency", *Adv. Electromag.*, vol. 5, no. 3, pp. 42–49, 2016 (DOI: 10.7716/aem.v5i3.425).

[15] M. Drapalik, J. Schmid, E. Kancsar, V. Schlosser, and G. Klinger, "A study of the antenna effect of photovoltaic modules", *Int. Conf. on Renewable Energ. and Power Quality (ICREPQ'10)*, Granada, Spain, pp. 371–375, 2010 (DOI: 10.24084/REPQJ08.333).

[16] Standard EN 61000-6-3, "Electromagnetic compatibility (EMC) – Part 6–3: Generic standards – Emission standard for residential, commercial and light-industrial environments", 2008/A1:2012 [Online]. Available: https://webstore.iec.ch/publication/27413

[17] Standard EN 61000-6-4, "Generic standards – Emission standard for industrial environments", 2019 [Online]. Available: https://webstore.iec.ch/publication/26622

[18] S. Bartnikowska, A. Olszewska, and W. Czekała, "The current state of connection issues of renewable energy sources installations to the electrical grid", *Energy Policy J.*, vol. 20, no. 2, pp. 117–128, 2017 [Online]. Available: https://epj.min-pan.krakow.pl/pdf-96172-28949?filename=The%20current%20state%20of.pdf (ISSN: 14296675) (in Polish).

[19] "How a photovoltaic installation works" [Online]. Available: https://2quantumenergy.com/functioning-of-pv-systems/

[20] K. Ranabhat *et al.*, "An introduction to solar cell technology", *J. of Applied Engineer. Sci.*, vol. 14, no. 405, pp. 481–491, 2016 (DOI: 10.5937/jaes14-10879).

[21] M. Arun, "Types of solar cells and its applications", *Int. J. of Sci. Devel. and Res.*, vol. 4, no. 2, pp. 260–267, 2019 [Online]. Available: https://www.ijsdr.org/papers/IJSDR1902043.pdf (ISSN: 24552631).

[22] W. Mao *et al.*, "A research on cascaded H-bridge module level photovoltaic inverter based on a switching modulation strategy", *Energies*, vol. 12, no. 10, 2019 (DOI: 10.3390/en12101851).

[23] A. B. Rifai and M. A. Hakami, "Health hazards of electromagnetic radiation", *J. of Biosciences and Medicines*, vol. 2, pp. 1–12, 2014 (DOI: 10.4236/jbm.2014.28001).

[24] Standard EN 50160, "Voltage characteristics in public distribution system", 2010 [Online]. Available: https://standards.iteh.ai/catalog/standards/clc/18a86a7c-e08e-405e-88cb-8a24e5fedde5/en-50160-2010

[25] Standard PN-IEC 60364, "Electrical installations in buildings", 2000 (in Polish).

[26] J. M. Flores-Arias *et al.*, "A memory-efficient true-RMS estimator in a limited-resources hardware", *Energies*, vol. 12, no. 9, 2019 (DOI: 10.3390/en12091699).

[27] K. Zhou and L. Cai, "An algorithm for calculating the RMS value of the non-sinusoidal current used in AC resistance spot welding", *J. of Power Electron.*, vol. 15, no. 4, pp. 1139–1147, 2015 (DOI: 10.6113/JPE.2015.15.4.1139).

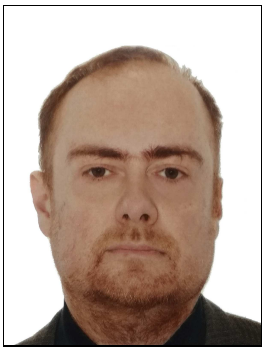
[28] Standard EN 62233, "Measurement methods for electromagnetic fields of household appliances and similar apparatus with regard to human exposure", 2008 [Online]. Available: https://standards.iteh.ai/catalog/standards/clc/a0bd973c-5a39-4cfa-8261-a474cd6a392f/en-62233-2008

[29] K. Maniak, "Badanie zjawisk elektromagnetycznych występujących na osuwiskach (The study of electromagnetic phenomena to landslides)", Doctoral Dissertation, Wrocław University of Technology, Wrocław, Poland, 2008 (in Polish).

[30] J. Miecznik, "Electromagnetic fields of dipole sources", Geologia, vol. 37, pp. 37–61, 2011 [Online]. Available: http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.baztech-article-AGHL-0002-0002/c/Miecznik.pdf (in Polish).

[31] A. Choroszucho and B. Butrylo, "Local attenuation of electromagnetic field generated by wireless communication system inside the building", *Przegląd Elektrotechniczny*, vol. 87, pp. 123–127, 2011 [Online]. Available: http://pe.org.pl/articles/2011/7/31.pdf

[32] I. Kolcunová *et al.*, "Infuence of electromagnetic shield on the high frequency electromagnetic field penetration through the building material", *Acta Physica Polonica*, vol. 131, pp. 1135–1137, 2017 (DOI: 10.12693/APhysPolA.131.1135).



Krzysztof Maniak received his M.Sc. and Ph.D. degrees in Electronics from Wroclaw University of Technology, Poland, in 2001 and 2008, respectively. Since 2012, he has been with the National Institute of Telecommunications in Wrocław. He is an author and co-author of numerous research papers and publications. His professional work focuses on: power systems, electromagnetic compatibility and design of measuring equipment. He also participates in research and measurement-related activities. His research interests focus on the testing of power grid quality and its broadband interference protection.

https://orcid.org/0000-0003-1974-909X
E-mail: k.maniak@il-pib.pl
National Institute of Telecommunications
Szachowa 1
Warsaw, Poland



Tomasz Tomczyk received his M.Sc. degree from the Faculty of Electronics, Wrocław University of Technology, in 2005. Since 2012, he has been with the National Institute of Telecommunications in Wrocław. He works at the EMC research laboratory. His interests cover broadly defined EMC research topics, especially in-situ testing.

https://orcid.org/0000-0003-2870-9241
E-mail: t.tomczyk@il-pib.pl
National Institute of Telecommunications
Szachowa 1
Warsaw, Poland



Karolina Spalt received her M.Sc. degree in Electronics from Wroclaw University of Technology, Poland, in 2006. Since 2006, she has been with the National Institute of Telecommunications in Wrocław. She is an author and co-author of numerous research papers and publications. Her professional work focuses on: electromagnetic compatibility, calibration of measurement equipment and measurement of electromagnetic fields. She is the head of the Laboratory of EMC Measuring Apparatus and participates in research and measurement-related activities.
E-mail: k.spalt@il-pib.pl

National Institute of Telecommunications
Szachowa 1
Warsaw, Poland



Janusz Sobolewski received his Ph.D. degree in Electronics from Wrocław University of Technology, Poland, in 1980. Since 1980, he has been an Assistant Professor with the National Institute of Telecommunications in Wrocław and the head of the EMC Department at NIT Wrocław (2005–2019). His research interests include frequency management and frequency coordination (national and international), radio network analysis, planning and optimization, as well as electromagnetic compatibility of radio networks.

https://orcid.org/0000-0001-5816-8461
E-mail: j.sobolewski@il-pib.pl
National Institute of Telecommunications
Szachowa 1
Warsaw, Poland



Jacek W. Wroński received his Ph.D. degree from Wroclaw University of Technology (Poland). He is the Chief Research and Scientific Specialist at the Spectrum Management and Engineering Section of the EMC Department of the National Institute of Telecommunications, National Research Institute, Poland. Over the period of more than 25 years, he has developed numerous software solutions and tools used in radiocommunications. He is interested in programming-related aspects, especially in efficiency and accuracy of optimization, stability of numerical method, multithreading, massive parallel computation and remoting. He is the author or co-author more than 120 NIT Reports (scientific and commercial), as well as over 20 conference papers and publications.

https://orcid.org/0000-0002-0452-3541
E-mail: j.wronski@il-pib.pl
National Institute of Telecommunications
Szachowa 1
Warsaw, Poland

Paper

Impact of Structures on the Operation of Air Traffic Radiolocation and Radionavigation Surveillance Systems

Maciej J. Grzybkowski, Daniel Niewiadomski, and Marcin Mora

National Institute of Telecommunications, Warsaw, Poland

https://doi.org/10.26636/jtit.2021.155221

Abstract—The article describes interference affecting the operation of radiolocation and radionavigation devices used in the air traffic surveillance systems, caused by the proximity of building structures. The impact of a hypothetical structure on the operation of primary and secondary air traffic surveillance radars and DVOR/DME beacons was simulated. The results of this simulation are presented in the form of airspace sectors in which false identification of aircraft may occur, and where it will not be possible to identify aircraft or use beacons due to the certain portion of airspace being in the shadow created by the structure. Analysis of the possibility of the PSR radar receiver being blocked by a strong signal reflected from a nearby building was performed as well.

Keywords—air traffic safety, environment obstacles, ground radiolocalization systems.

1. Introduction

Civilian (and military) aviation authorities aim to ensure the safety of personnel and passengers traveling by air. The desired level of safety is ensured by complying with organizational rules and meeting applicable technical requirements. Ground-based air traffic surveillance systems are an important aspect of technical flight safety measures. Air traffic control systems monitor aircraft traveling enroute, during the departure and approach phase of the flight (around airports) and at airports themselves. Air traffic surveillance is performed with the use of various radio-electronic solutions, such as radionavigation, radiolocation and air radiocommunication systems. Ground-based air traffic surveillance hardware is usually located at airport premises, near the runways, but may be also positioned at other locations of key significance for air traffic control. Such hardware interacts with devices installed aboard the aircraft.

Terrestrial radionavigation systems ecomprise, inter alia, non-directional beacons (NDB), distance measuring equipment (DME), VHF Omni Directional Range (VOR), and VOR using the Doppler effect (DVOR) azimuth beacons, markers (route beacons) and Instrument Landing System(s) (ILS) – solutions made up of directional beacons that assist crews while landing in low visibility conditions, or their microwave versions, known as Microwave Landing System (MLS). Ground-based radiolocation systems use radars to indicate the position of the aircraft in airspace, they are primary surveillance radars (PSRs), and to identify of aircraft and provide much more information on it, they are secondary surveillance radars (SSRs) [1].

Correct operation of air traffic surveillance systems is crucial and their performance should not be affected by any potential disturbances. The sources of such disturbances may include electromagnetic (active) sources, or emission caused by environmental conditions (passive). Specifications protecting radionavigation and radiolocation systems against interference originating from various sources are provided for in numerous documents published, for instance, by the International Telecommunications Union [2]–[6]. However, disturbance affecting the operation of ground-based air traffic surveillance devices, caused by closely located structures, including various types of buildings, is a problem as well. In some publications, e.g. [7]–[9], the authors noted the potential impact that man-made obstacles, such as buildings, may exert on the operation of radars. However, they failed to present any methods for analyzing the outcomes of such impact.

Structures present in close proximity of navigation systems need to be positioned in a manner allowing to minimize their negative impact. Some typical types of interference include electromagnetic wave reflections and shadowing effects rendering specific portions of airspace unsuitable for monitoring. Therefore, restrictions concerning the height of buildings existing in close proximity to these radio devices are introduced. As it is impossible to do away with all structures in the vicinity of the abovementioned devices (e.g. airport terminals, hangars, air traffic control towers, or even offices and hotels), it is recommended that their impact on the operation of such systems be studied in order to gather information about the potential adverse types of impact.

This article covers the effects of passive impact that structures may exert on the operation of air traffic surveillance systems. A scenario in which fictitious structures impact the operation of hypothetical radiolocation devices located within the airport perimeter will be simulated as well, and an analysis of the effects of such impact will be performed.

In the literature, e.g. in [10]–[12], theoretical considerations related to reflections of waves generated by SSR radars, caused by ground-based objects, as well as to the scattering of radio waves by wind turbines may be found. However, it is the practical effects of these reflections, as well as the outcomes of radio wave scattering and shadowing phenomena that are extremely important and interesting for the aviation authorities.

2. Effects of Disturbance Affecting Air Traffic Surveillance Systems

False aircraft locations caused by the reflection of electromagnetic waves are one of the most serious effects of disturbance created by structures present in the vicinity of terrestrial radionavigation or radiolocation equipment. In addition to such false locations, coordinate- and altitude-related errors in determining the azimuth of and the distance to beacons may be experienced as well. These phenomena are caused by the multi-path propagation of the radio wave in scenarios in which other objects (structures), reflecting the emitted signals are located in the vicinity of the tracking facilities or the tracked object.

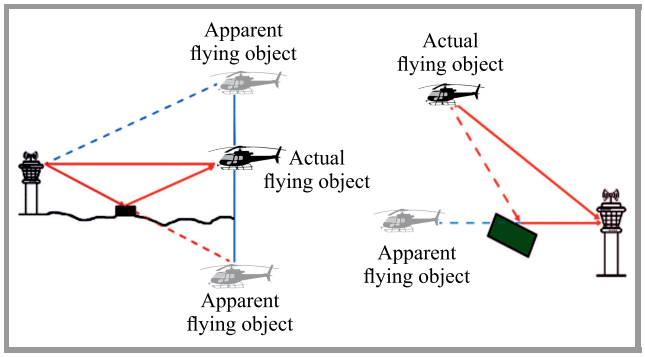


Fig. 1. False location of a flying object identified when the radio wave is reflected by a building [1].

Such cases are illustrated in Fig. 1. Radio shadowing caused by structures is another adverse effect. In this scenario, certain space exists that cannot be reached by radio waves emitted from the transmitters of radiolocation devices, as it is located below the line of sight. The radio shadowing phenomenon is somewhat alleviated by tropospheric refraction curving the trajectory of radio waves. Antennas of radiolocation and radio navigation devices are designed in such a way that their vertical radiation pattern covers a large range of elevation angles in relation to the Earth's surface, i.e. both small angles for low-flying and distant aircraft and large angles for aircraft flying high and close to the antenna. However, the resulting radio shadowing effect limits radar or beacon coverage, as radio waves fail to reach flying objects located in a certain part of the airspace. On the other hand, from the point of view of aircraft, objects located in the proximity of ground-based radio navigation systems weaken their signals which cannot be received, in some cases, by planes. Due to high operating frequency used, the diffraction of radio waves observed at the edges of buildings is small, but the radio wave is almost completely suppressed in the radio shadow zone. Therefore, obstruction zones are defined, in radiolocation, both in terms of their horizontal and vertical planes. These sectors are the portions of airspace in which aircraft are undetectable or where ground equipment is undetectable by aircraft.

Contact with ground-based radiolocation systems may also be lost due to radar receiver blocking (saturation). This may be the case in a situation in which a high-power radar pulse is generated and directed towards objects located nearby. The signal reflected by these objects returns into the aperture of the radar antenna. Because of the high power of the reflected pulse signal, a short break in the receiver's operation may occur due to the saturation of its front-end circuit. This phenomenon is quite short-lasting due to the duration of the reflected signal pulse and the rotation speed of the radar antenna. But even such short breaks in the operation of the receiver make it impossible to distinguish objects located further away, thus temporarily preventing objects positioned at greater distances from being identified. The receiver blocking phenomenon applies only to PSR radars in which the emission and reception of pulsed signals occurs at the same frequency. SSR radar receivers are not blocked by reflections from nearby objects, as they operate based on a different principle. Such radars use two different frequencies for transmitting and receiving.

To avoid blocking a PSR radar, the power of the signal arriving at its receiver, after being reflected from a (nearby) structure P_o, should be lower than the receiver's saturation power P_{bl}.

$$P_o \leq P_{bl} \ . \tag{1}$$

P_o may be determined by:

$$P_o = P_{pr} + G_r(\Theta_r) - L_b - L_{fo} + M - R \ \text{[dBm]} \quad , \tag{2}$$

where:

P_{pr} – equivalent power radiated by the radar isotropically (in a pulse) [dBm],

$G_r(\Theta_r)$ – energy gain of the receiving antenna in relation to the isotropic antenna [dBi],

Θ_r – antenna elevation angle considering the slope of the antenna pattern,

L_b – attenuation of the radio path on the incident wave and reflected from the object towards the antenna [dB],

L_{fo} – receiver feeder attenuation [dB],

M – margin factor considering multipath propagation [dB],

R – object reflection coefficient [dB].

In most cases, the M factor may be left out because the receiver is always blocked when the reflecting objects are very close to the radar (are located in the near field zone). Therefore, it is assumed that the object from which the radio wave is reflected, being capable of blocking the receiver, is not located in the near field zone of the PSR radar antenna. The worst conditions for blocking the radar receiver occur when waves propagate in free space, without scattering, atmospheric absorption and when the object's reflection

coefficient equals 0 dB. When needed, different reflection coefficient values for various frequencies and wave polarizations, as well as for different types of building materials may be calculated according to the methodology presented in [13].

It should be emphasized that modern PSR radars are equipped with a sensitivity time control (STC) device used to suppress strong signals generated by reflections at very short distances. Nevertheless, any structure present in close proximity to the radar may reflect radio waves. Hence, an analysis of a scenario in which the receiver may be potentially blocked by reflections should always be performed.

3. Simulation of Interferences Caused by Structures

Simulations determining the impact of nearby buildings on the operation of ground air traffic control devices should allow to designate radio-limited zones. When designing structures present in the vicinity of radiolocation devices, such an analysis should be carried out in order to obtain information about potentially incorrect operation of equipment. In order to assess spatial limitations of radar (loss of coverage along a specific direction and at a specific height above ground level) caused by the formation of radio shadow zones, the solid of the planned structure should be modeled as a terrain obstacle. Next, the elevation angles and azimuths of the beam radiated by the radar incident on the analyzed structure should be determined. Such a procedure allows to identify areas in which radio shadows caused by this structure are formed, i.e. areas in which radio visibility is reduced. Similarly, areas with reflections that may result in false aircraft locations being determined may be identified.

In order to simulate disturbances affecting the operation of radiolocation devices and caused by ground structures, a fictitious building, i.e. hotel complex, was modeled in the vicinity of one of the inactive airports in Lower Silesia, Poland. Fictitious PSR and SSR radars, as well as DVOR/DME beacons, were modeled, too. Their layout is shown in Fig. 2. Calculations related to airspace zones in which radio shadows may appear and in which locations of aircraft may be determined falsely relied upon a detailed digital map (digital terrain model – DTM).

Tables 1–3 present selected basic parameters of typical PSR and SSR radars, as well as of DVOR/DME beacons used for the purpose of the simulation described in this article.

Table 1
Selected basic parameters of a PSR radar

Operating frequency [MHz]	2800
Mechanical antenna tilt [°]	0
Antenna gain – high beam [dBi]	32
Antenna gain – low beam [dBi]	33
Pulse radiation power [kW]	32
Height of the electrical center of the antenna [m] a.g.l.	37

Table 2
Selected basic parameters of the SSR radar

Operating frequency [MHz]	1030/1090
Mechanical antenna tilt [°]	0
Height of the electrical center of the antenna [m] a.g.l.	30

Table 3
Selected basic parameters of the DVOR/DME

Type of DVOR antenna	Alford slot antenna
Type of DME antenna	Omnidirectional
Height of the electrical center of the DVOR antenna [m] a.g.l.	4.7
Height of the electrical center of the DME antenna [m] a.g.l.	8

3.1. Simulation of the Shape of the Radio Shadow Zone Caused by Hotel Buildings

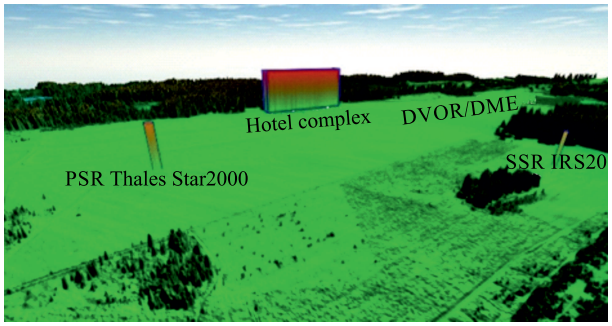


Fig. 2. Ground-based aviation equipment and hotel buildings modeled on the DTM map base.

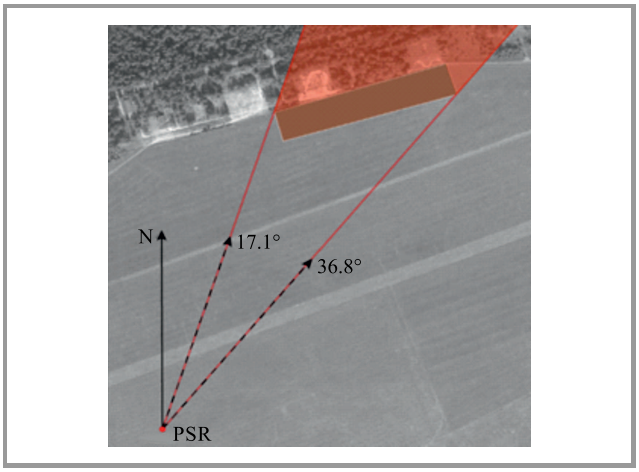


Fig. 3. Azimuth range in which the horizontal diaphragm/radio shadowing will occur for PSR radar.

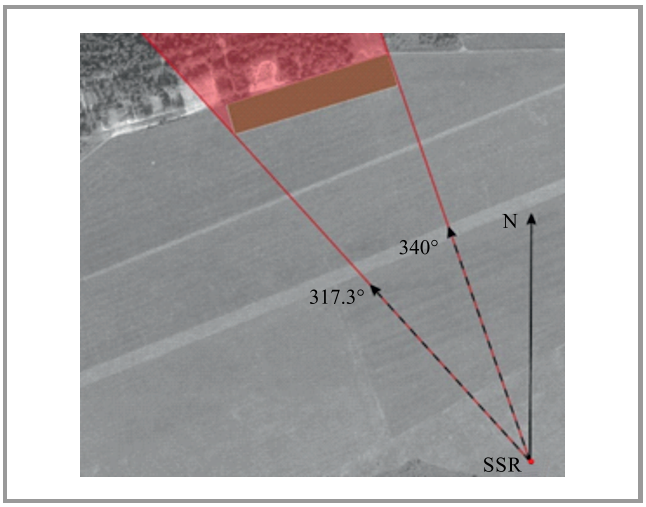


Fig. 4. Azimuth range in which the horizontal diaphragm/radio shadowing will occur for SSR radar.

By simulating the building's impact on the operation of PSR and SSR radars, a radio shadow zone, horizontal and vertical diaphragms of PSR and SSR radars, as well as areas with potential range losses could be determined. Figures 3 and 4 show the calculated azimuth ranges with the horizontal diaphragm of PSR and SSR radars.
Figures 5 and 6 show the calculated azimuth ranges in which the vertical diaphragm of PSR and SSR radars will occur. The following designators are used in these drawings:

- Φ – angle of deviation from the direction of the maximum radiation,

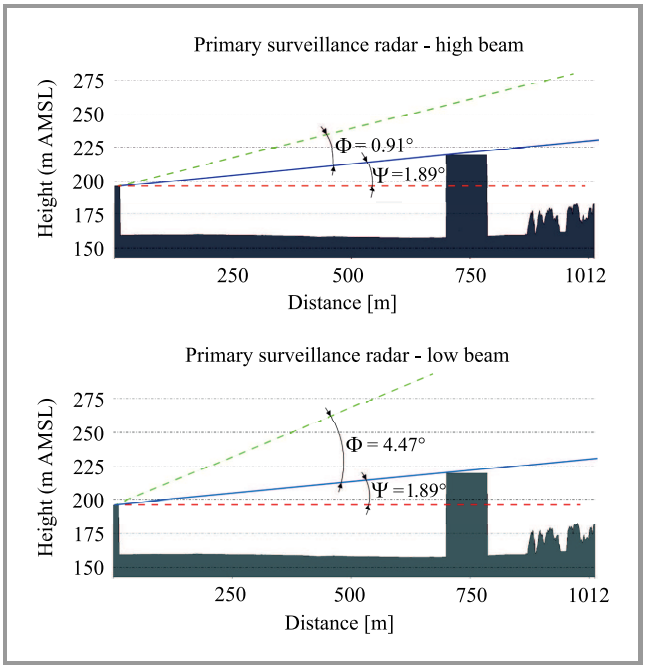


Fig. 5. Azimuth range in which the vertical diaphragm/radio shadowing will occur for PSR radar. (see the digital edition for color images)

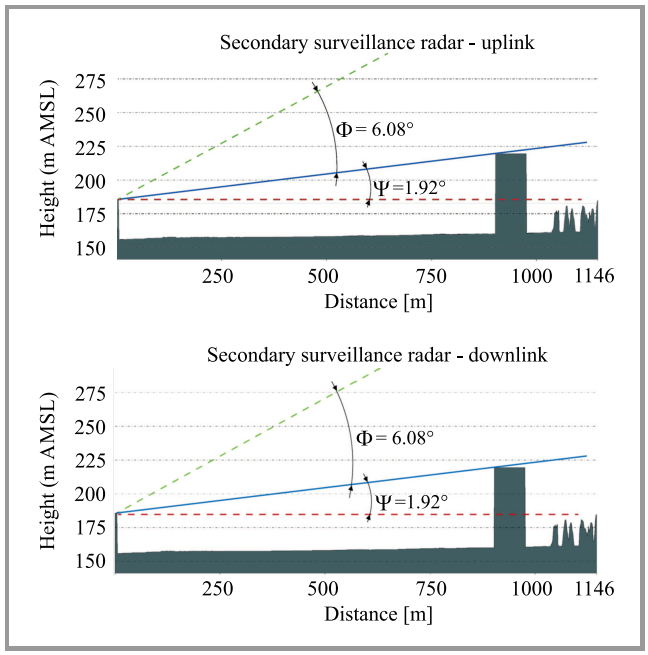


Fig. 6. Azimuth range in which the vertical diaphragm/radio shadowing will occur for SSR radar.

- Ψ – angle of the building's facade,
- dashed red line – height of the electrical center of the antenna above ground level,
- dashed green line – direction of maximum radiation,
- solid blue line – radar antenna's direct line of sight.

Using the radiation characteristics of the radar antenna system, the areas of the airspace (with height being a function of distance) in which radar coverage is lost have been determined. These areas were designated for both PSR radiation beams, i.e. for the upper and lower beam, and for the SSR radar radiation beams serving the downlink and uplink.

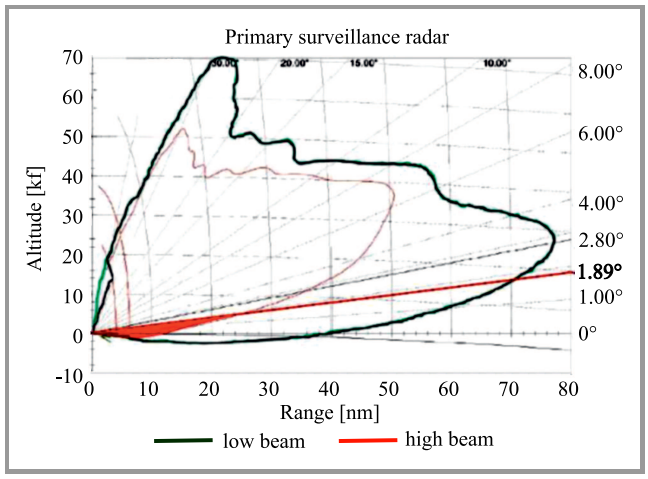


Fig. 7. The area (marked red) in which PSR radar coverage is lost due to the screening effect caused by the hotel building (radar high beam).

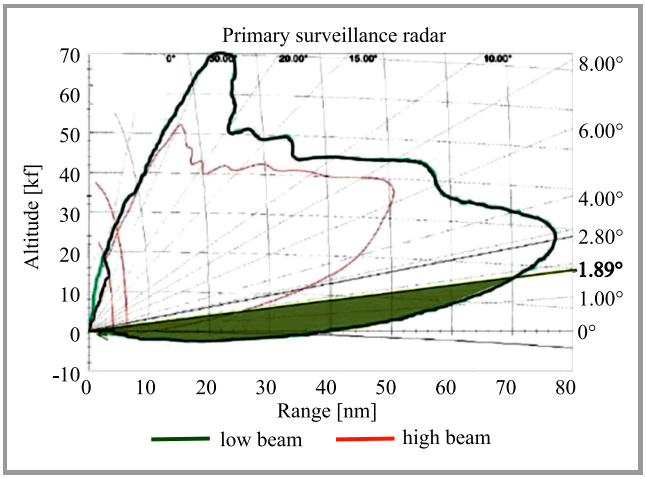


Fig. 8. The area (marked green) in which PSR radar coverage is lost due to the screening effect caused by the hotel building (radar high beam).

Figures 7 and 8 show the area with PSR radar coverage lost, while Figs. 9 and 10 show the area with SSR radar coverage lost.

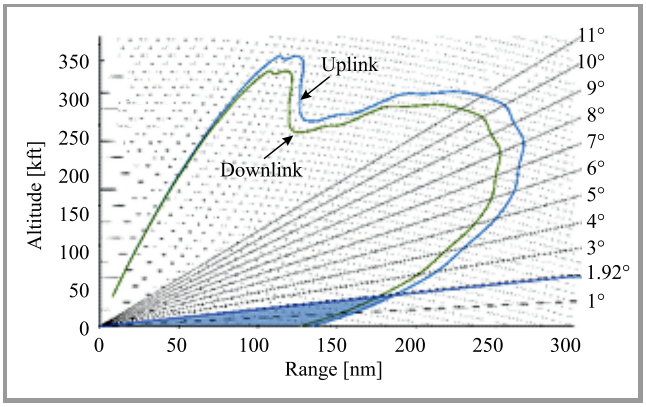


Fig. 9. The area (marked blue) in which SSR radar coverage is lost due to the screening effect caused by the hotel building (beam serving the uplink).

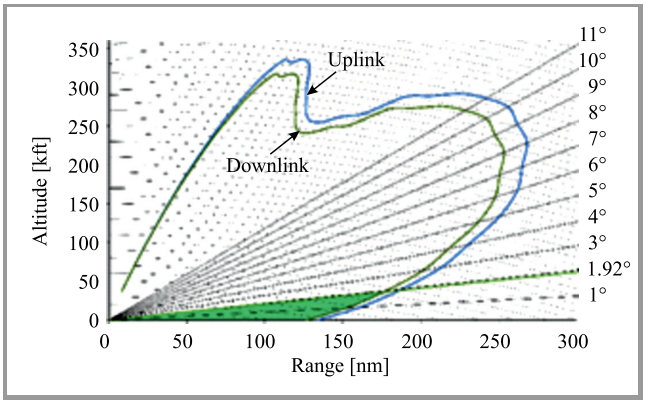


Fig. 10. The area (marked green) in which the SSR radar coverage is lost due to the screening effect caused by the hotel building (beam serving the downlink).

Table 4 presents the boundaries of the space in which loss of the PSR radar coverage is expected (with height above ground level being a function of distance), occurring within the azimuth range of 17.1–36.8°.

Table 4
Loss of PSR radar coverage

High beam	Radar operating range [nautical miles]	10	15	20
	Radar operating ceiling [thousands of feet]	0.2	0.3	0.4
Low beam	Radar operating range [nautical miles]	20	40	70
	Radar operating ceiling [thousands of feet]	0.4	0.8	1.6

Table 5 shows the boundaries of the space in which loss of the SSR radar coverage is expected (with height above ground level being a function of distance), occurring within the azimuth range of 317.3–340°.

Table 5
Area with PSR radar coverage lost

Uplink beam	Radio operating range [nautical miles]	50	100	180
	Radar operating ceiling [thousands of feet]	13	30	70
Downlink beam	Radar operating range [nautical miles]	50	100	180
	Radar operating ceiling [thousands of feet]	12	30	60

3.2. Identification of Zones with Potential False Detection of Flying Objects

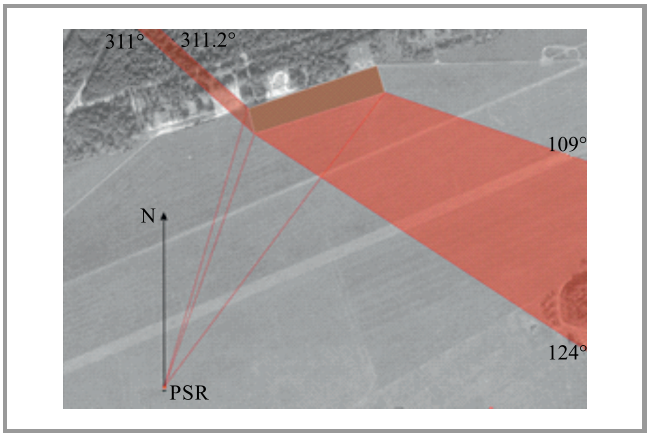


Fig. 11. Calculated azimuths at which primary radar beam will reflect from the hotel buildings in the horizontal plane.

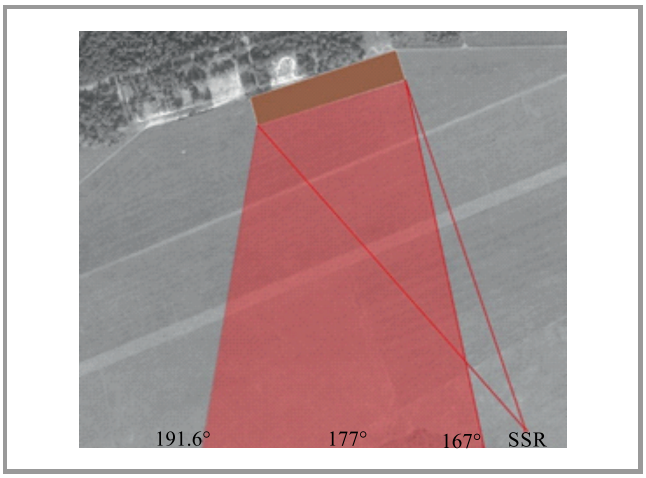


Fig. 12. Calculated azimuths values at which secondary radar beam will reflect from the hotel buildings in the horizontal plane.

In the scenario in which radar beams may be reflected by the hotel buildings, thus leading to false determination of the position of flying objects, calculations of horizontal and vertical plane azimuth ranges were performed. Figures 11 and 12 show the ranges of radar beam reflections from hotel buildings in the horizontal plane, for PSR and SSR radars. In the case of primary radar, a location error may occur in the azimuth ranges of 109–124° and 311–311.2°.

For secondary radar, incorrect detection due to radar beam may occur in the azimuth range of 167–191.6°.

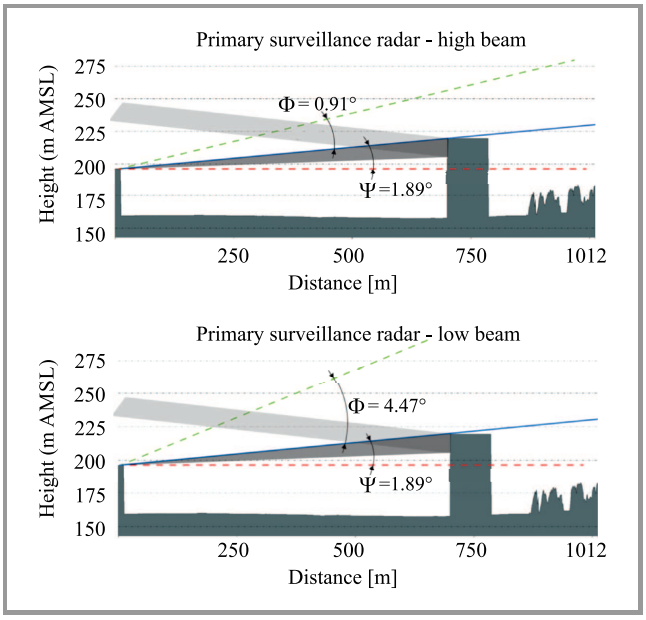


Fig. 13. Calculated azimuth range at which primary radar beam will reflect from the hotel buildings in the vertical plane.

Figures 13 and 14 show the range of radar beam reflections from hotel buildings in the vertical plane, for PSR and SSR radar, respectively. The markings of angles and lines are the same as in Figs. 5 and 6.

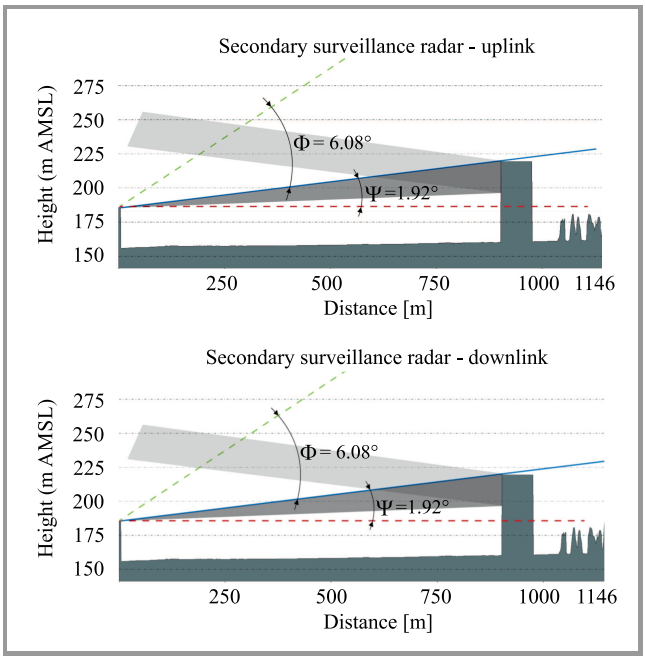


Fig. 14. Calculated azimuth range at which secondary radar beam will reflect from the hotel buildings in the vertical plane.

3.3. Simulation of Disturbances Affecting the Operation of DVOR/DME Beacons

When considering limitations affecting the operation of a beacon, an analysis of its location in relation to buildings should be performed in order to determine the form of the diaphragm/radio shadowing. It causes loss of the beacon's operating range. Apart from being diffracted at the edges of the structure, it was assumed that the radio wave is completely suppressed in the shadow zone, making it impossible to identify the beacon of flying objects. In addition, it is necessary to establish the boundaries of the airspace in which false detections of the beacon's position may take place due to reflections from the walls of the building.

Results of analyses concerned with the existence of the radio shadowing zone are presented in the Figs. 15–16. Figure 15 shows the determined loss of operating range of

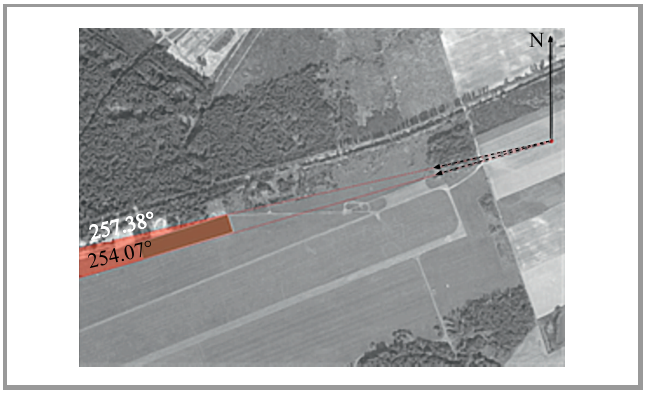


Fig. 15. Calculated azimuths range at which the horizontal diaphragm of DVOR/DME beacons will occur.

DVOR/DME beacons due to the diaphragm/radio shadowing caused by the hotel buildings in the horizontal plane. The loss of beacon's range (no identification possible) may occur within the azimuth range of 254.07–257.38°.

Figure 16 shows the determined loss of operating range of DVOR/DME beacons due to the diaphragm/radio shadowing caused by the hotel buildings in the vertical plane.

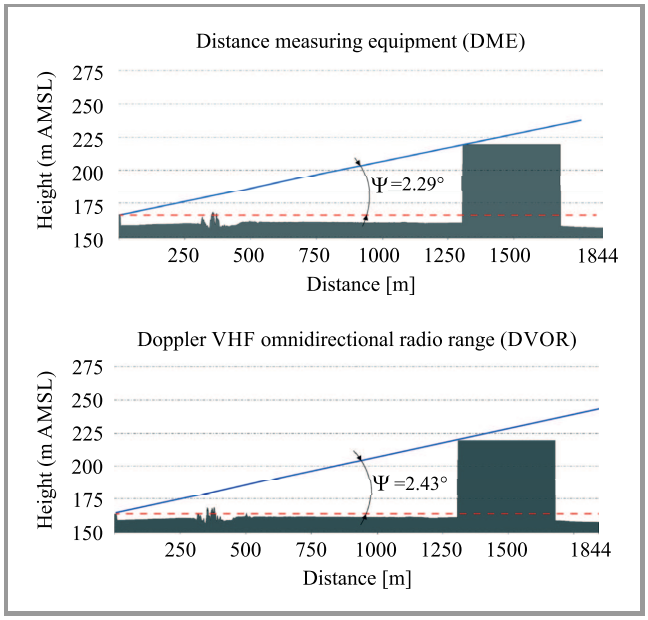


Fig. 16. Vertical azimuths below which signals of DVOR/DME beacons are obscured.

It was determined that in this case the loss of the ability to identify the beacon in the vertical plane, due radio shadowing, will be experienced below the elevation angle of 2.29° and for 2.43° for DME and DVOR, respectively.

Results of the analysis allowing to delineate the portion of the airspace in which flying objects may erroneously determine the location of beacons due to the reflections from the building, are shown in Figs. 17 and 18.

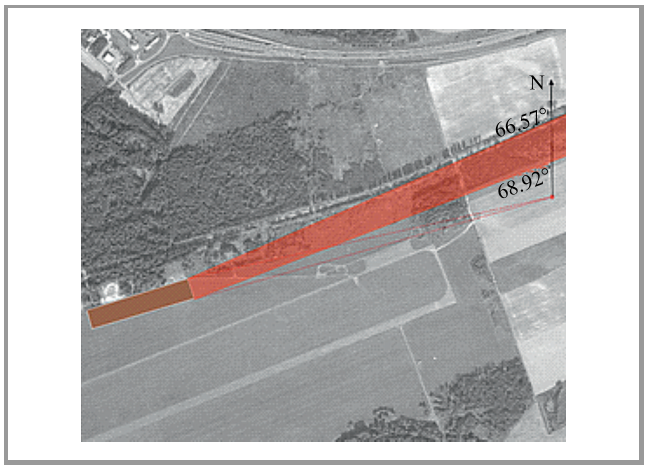


Fig. 17. Azimuth range within which horizontal reflections of the beams generated by DVOR/DME beacons from hotel buildings will occur.

Figure 17 illustrates, in the horizontal plane, the range of reflection of the signal beam from DVOR/DME beacons. In this case, the azimuth range of the reflected beacon beam is: 66.57–68.92°.

Finally, Fig. 18 presents the range of reflection of the beam generated by DVOR/DME beacons, from hotel buildings, in the vertical plane.

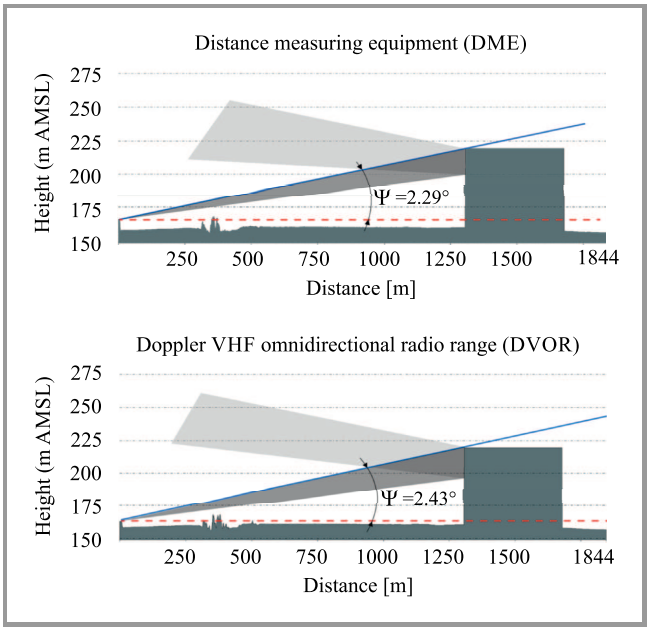


Fig. 18. Vertical azimuths below which beams of radio waves generated by DVOR/DME beacons are reflected by the building.

In this case, reflections of the radio waves generated by the beacons in the vertical plane, resulting from the presence of an obstacle in the form of a hotel building, will occur below the elevation angles of 2.29° and 2.43° for DME and DVOR, respectively. It means that low-flying aircraft may falsely determine the location of these beacons while traveling within such a range.

3.4. Analysis of Scenarios in which the PSR Radar Receiver is Potentially Blocked

Here, an analysis of the scenario in which the PSR radar circuit is potentially blocked upon receiving a strong signal reflected from nearby objects is performed with the worst propagation conditions taken into consideration and assuming that the beam radiated by the radar is fully reflected from these objects.

In this case, the radio path loss for the signal emitted by the radar and reflected from the object can be determined by the well-known formula for free-space basic transmission loss [16]:

$$L_b = L_{bf} = 32.4 + 20\log f + 20\log d \text{ [dB]} , \qquad (3)$$

where f is the frequency in [MHz] and d is the distance in [km].

Knowing the distance between the PSR radar and the nearby object, and being aware of the height of the radar antenna's electric center and the height of the object concerned, one may assess whether the radar receiver will be blocked by the reflection of the radiated beam.

Assuming the typical parameters of a PSR radar: $P_{pr} = 32$ kW $= 75.1$ dBm, $f = 2800$ MHz, $G_r(\theta_r) = 33.0$ dBi and $L_{fo} = 2$ dB, and using the distance from the radar antenna to the nearest building of approx. 693 m, the loss of free-space path from the radar to the object is:

$$L_{bf} = 32.4 + 20\log 2800 + 20\log(2 \times 0.693) = 32.4 + 68.9 + 2.8 = 104.1 \text{ dB} . \quad (4)$$

Using the relative decrease in the gain of the receiving antenna, observed as the distance towards the object decreases (calculated based on the vertical characteristics of the PSR radar), for an incidence angle of 0° in relation to the azimuth of maximum radiation (3 dB), the total gain of this antenna towards the considered building will be $33.0 - 3 = 30.0$ dB. With the assumed reflection coefficient of $R = 1$ (0 dB), the power of the signal received by the radar after reflection by the building is:

$$P_o = 75.1 + 30.0 - 104.1 - 2 - 0 = -1.0 \text{ dBm} . \quad (5)$$

According to [4], the average saturation level of an air traffic control (ATC) radar receiver is equal to $P_{bl} = 13$ dBm. Thus, according to Eq. (2), the received signal power obtained above, after reflection from a (near) object P_o is by 14 dB lower than the saturation power of the radar's receiver P_{bl}. In the scenario under consideration, the building will not block the operation of the PSR radar.

4. Structure-related Restrictions Imposed by ICAO

In order to reduce the probability of disturbances affecting the operation of ground-based air navigation equipment, caused by the close proximity of various building structures and radio shadowing, the International Civil Aviation Organization (ICAO), has defined the recommended minimum distances of radionavigation, radiolocation and air communication devices from nearby buildings [14]. The methodology relied upon for determining the boundaries of off-limits areas, i.e. the sizes of protection zones around ground-based air navigation equipment, is presented in Fig. 19, while the distances applicable to various types of such devices are shown in Table 6. It should be noted that apart from various types of structures, no other objects (mobile and stationary, permanent or temporary), terrain faults or vegetation of any kind should be present in the protected zones. If a structure needs to be positioned inside the protection zone, it is necessary to analyze its impact on the operation of each of the ground-based devices in operation.

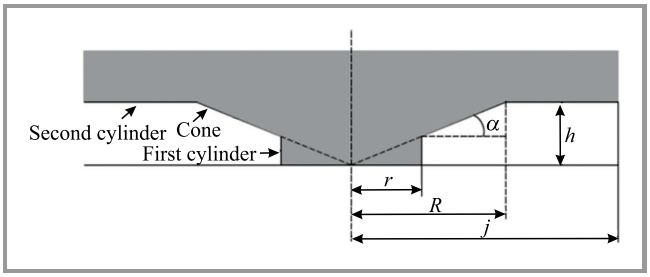


Fig. 19. Protection zones for aviation ground equipment [14].

Some general guidelines (but not having the form of formal standards) regarding the presence of buildings in the vicinity of DVOR/DME devices are also presented in [15].

ICAO recommends to designate protection zones around radiolocation devices and to define restricted areas in which the no structures should be present. These recommendations should be obeyed into practice. However, in some cases, when it is not possible to avoid the construction of such structures, it is necessary to identify those sectors of the airspace in which proper operation of radiolocation equipment may be affected.

Table 6
Data from Fig. 19 for selected radionavigation and radiolocation devices

Type of navigation or surveillance facilities	First cylinder radius r [m]	Cone wall inclination angle a [°]	Second cylinder radius R [m]	Second cylinder radius* j [m]	Second cylinder height* r [m]	Origin of cone and axis of cylinders at ground level
Distance measuring equipment DME N	300	1.0	3000	n/a	n/a	Base of antenna
Doppler VHF omnidirectional range DVOR	600	1.0	3000	10000	52	Centre of antenna systems
Primary surveillance radar PSR	500	0.25	15000	n/a	n/a	Base of antenna
Secondary surveillance radar SSR	500	0.25	15000	n/a	n/a	Base of antenna
* for wind turbines only						

5. Conclusions

Data sourced from ground-based radiolocation and radionavigation systems are mainly used to identify aircraft, to determine their current location in the airspace and some of their flight-related parameters. The position of radionavigation devices on the ground is determined by the aircraft as well. It may be difficult or outright impossible to obtain such information when building structures and other man-made objects are located in the vicinity of air traffic control stations. In extreme cases, as a result of reflections of radio waves radiated by radar antennas, false positions of the aircraft may be presented. On the other hand, the radio shadowing phenomenon caused by the existence of terrain, artificial and natural obstacles, results in the operation of radiolocation devices being subjected to some spatial limitations.

The effects of such impact are presented based on the example of a fictitious building located in the vicinity of a runway of a closed airport in Poland. Those sectors of the airspace in which the positions of aircraft may be determined falsely or cannot be determined altogether due to radio shadowing have been established. The probability of blocking the radar receiver by radio waves it generates being reflected from a nearby object, was analyzed as well. In addition, those airspace sectors have been identified where it would be impossible for aircraft to identify the position of radio beacons due to the formation of dead zones (radio shadowing) caused by the presence of a building.

The methods relied upon for analyzing the impact of surroundings on the operation of radiolocation and radionavigation systems, as presented in this article, may facilitate forecasting unfavorable phenomena adversely affecting the operation of these systems (in a scenario in which buildings already exist in the vicinity of airports) and may also contribute to eliminating them at the design stage.

The calculation results presented indicate that the dimensions of their projection on the plane perpendicular to the direction of the radar devices were as small as possible (surface area and height).

In such scenario, the radio shadow created by the building will be minimized. In addition, the value of the angle at which false readouts indicating the position of flying objects or beacons will be minimized as well.

References

[1] M. J. Grzybkowski, "Environmental restrictions of the operation of air traffic radiolocation and radionavigation systems", *KKRRiT*, Warsaw, Poland, 2019, pp. 362–366 (DOI: 10.15199/59.2019.6.51).

[2] Recommendation ITU-R M.1462-1, "Characteristics of and protection criteria for radars operating in the radiolocation service in the frequency range 420–450 MHz", 2019 [Online]. Available: https://www.itu.int/rec/R-REC-M.1462/

[3] Recommendation ITU-R M.1463-3, "Characteristics of and protection criteria for radars operating in the radiodetermination service in the frequency band 1215–1400 MHz", 2015 [Online]. Available: https://www.itu.int/rec/R-REC-M.1463/

[4] Recommendation ITU-R M.1464-2, "Characteristics of non-meteorological radiolocation radars, and characteristics and protection criteria for sharing studies for aeronautical radionavigation and radars in the radiodetermination service operating in the frequency band 2700–2900 MHz", 2015 [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.1464-2-201502-I!!PDF-E.pdf

[5] Recommendation ITU-R M.1638-1, "Characteristics of and protection criteria for sharing studies for radiolocation (except ground based meteorological radars) and aeronautical radionavigation radars operating in the frequency bands between 5250 and 5850 MHz", 2015 [Online]. Available: https://www.itu.int/rec/R-REC-M.1638-1-201501-I/

[6] Recommendation ITU-R M.1461-2, "Procedures for determining the potential for interference between radars operating in the radiodetermination service and systems in other services", 2015 [Online]. Available: https://www.itu.int/rec/R-REC-M.1461-2-201801-I

[7] K. Frolic, "Developing buildings", 2020 [Online]. Available: https://www.pagerpower.com/news/planning-issues-for-buildings-radar-interference/

[8] S. T. Shipley *et al.*, "GIS tools for radar siting and analysis", *22nd Int. Conf. on Interactive Informat. Process. Systems for Meteorology, Oceanography, and Hydrology*, Atlanta, GA, USA, 2006 [Online]. Available: https://ams.confex.com/ams/Annual2006/techprogram/paper_104269.htm

[9] S. Shipley, D. Berkowitz, and R. M. Steadham, "Comparison of virtual globe technologies for depiction of radar beam propagation effects and impacts", *AGU Fall Meeting Abstracts*, San Francisco, CA, USA, 2007 [Online]. Available: https://ams.confex.com/ams/pdfpapers/135325.pdf

[10] G. Greving, "Numerical simulations of environmental distortions by scattering of objects for the radar – SSR and Flat Roofs, RCS and Wind-turbines", in *Proc. of the 3rd European Radar Conf.*, Machester, UK, 2006, pp. 241–244 (DOI: 10.1109/EURAD.2006.280319).

[11] G. Greving, W.-D. Biermann, and R. Mundt, "On the advanced simulation of distortions for navigation, landing and radar systems - modern methods, cases and results", *13th IAIN World Congress*, Stockholm, Sweden, 2009.

[12] G. Greving, W.-D. Biermann, and R. Mundt, "Dislocation aspects of wind-turbines within a wind-farm in some distance to primary radar, *18th Int. Radar Symp. IRS*, Prague, Czech Republic, 2017 (DOI: 10.23919/IRS.2017.8008170).

[13] Recommendation ITU-R P.2040-1, "Effects of building materials and structures on radiowave propagation above about 100 MHz", 2015 [Online]. Available: https://www.itu.int/rec/R-REC-P.2040-1-201507-I/

[14] ICAO, "European guidance material on managing building restricted areas, third edition", 2015 [Online]. Available: https://www.icao.int/EURNAT/EUR%20and%20NAT%20Documents/EUR%20Documents/EUR%20Documents/015%20-%20Building%20Restricted%20Areas

[15] ICAO, "Annex 10 to the Convention on International Civil Aviation, Aeronautical Telecommunications", *Radio Navigation Aids*, vol. 1, 2018 [Online]. Available: https://www.bazl.admin.ch/dam/bazl/en/dokumente/Fachleute/Regulationen_und_Grundlagen/icao-annex/icao_annex_10_aeronauticaltelecommunicationsvolumei-radionavigat.pdf.download.pdf/icao_annex_10_aeronauticaltelecommunicationsvolumei-radionavigat.pdf

[16] Recommendation ITU-R P.525-4, "Calculation of free-space attenuation", 2019 [Online]. Available: https://www.itu.int/rec/R-REC-P.525/en



Maciej J. Grzybkowski (Ph.D.) is currently working at NIT's EMC Department. His research areas cover wireless telecommunication network technologies, 5G and beyond 5G wireless networks, planning and optimization of radio networks, radio frequency management and spectrum engineering, as well as wireless telecommunications, electromagnetic compatibility and radiowave propagation modeling. He is highly experienced in regulations concerning the electronic communications sector and in assessing national telecommunication legislation: drafting opinions on current and future regulations for the telecommunications markets. He took part in many research and commercial projects acting in the capacity of a Project Manager, researcher or reviewer. He is an author or co-author over 50 publications and more than 100 research and commercial reports of the NIT.

https://orcid.org/0000-0001-8591-8627
E-mail: m.grzybkowski@il-pib.pl
National Institute of Telecommunications
Szachowa 1
Warsaw, Poland



Daniel Niewiadomski (M.Sc.) currently holds the position of a Research Engineer at Spectrum Engineering and Management Section, National Institute of Telecommunications, Poland. His main research areas cover planning and optimization of radio networks, radio wave propagation modeling, broadcasting networks, EMF exposure simulations and Internet of Things wireless networks. He is an author or co-author over 100 research and commercial reports published by NIT. He represents Poland at the following international regulatory forums: ITU-R Study Groups, ITU-R WRC, CEPT Working Groups and Project Teams.

https://orcid.org/0000-0002-5576-6551
E-mail: d.niewiadomski@il-pib.pl
National Institute of Telecommunications
Szachowa 1
Warsaw, Poland



Marcin Mora received his M.Sc. degree in Telecommunications from Wrocław University of Technology in 2017. Since 2017, he has been working as an engineering specialist at the Spectrum Engineering and Management Section, National Institute of Telecommunications. His areas of research interest include: planning and optimization of radio networks, EMF simulations, 5G Networks, digital terrestrial television, mobile broadband networks, digital audio broadcasting and IoT. He represents Poland at the following international regulatory forums: CEPT Working Groups and Project Teams. He is a co-author of 10 research and commercial reports published by NIT.

https://orcid.org/0000-0003-4427-6212
E-mail: m.mora@il-pib.pl
National Institute of Telecommunications
Szachowa 1
Warsaw, Poland

Paper

Accurate Location of Fiber Cable Fault with OTDR

Krzysztof Borzycki and Paweł Gajewski

National Institute of Telecommunications, Warsaw, Poland

https://doi.org/10.26636/jtit.2021.158621

Abstract—The paper reviews the factors limiting the accuracy of locating a fiber optic cable fault when using an optical time domain reflectometer (OTDR) and describes an error estimation method for typical use cases. The primary source of errors lies in the complex relationship between the length of the optical fiber (measured by OTDR), its routing, cable design depending on cable design and type of installation (i.e. duct, directly buried, aerial) as well as the spare lengths used for service purposes. The techniques which considerably improve the accuracy of the fault localization processes are presented, the importance of accurate documentation of the network and of referencing the fault location to the nearest splice instead of end of the line are discussed, as is the absence of cable helix factor in data sheets.

Keywords—fault location, fiber optic cable, helix factor, OTDR, single-mode fiber.

1. Introduction

OTDR is a measurement instrument used for diagnosing fiber optic networks [1], [2]. The main advantage of OTDR over a less expensive testing technique relying on a light source and a power meter is its ability to indicate the distance from the OTDR's optical port (or another reference location) to any discontinuity or "event" in the fiber being tested.

However, OTDR gives the length of the optical fiber between the OTDR's optical port and the fault. Due to the design and installation conditions, the measured length of the fiber is always longer than the length of the cable or the distance measured along its route. In order to be useful for fault location, such a value must be converted to cable route length to the fault or, much better, to the distance between the fault and the nearest component of the cable link that may be located without much difficulty – usually a cable splice.

Depending on the cable and installation technique used, the fiber length may exceed the route distance by up to 8%. With repeater sections in terrestrial fiber networks being up to 100 km long (or longer), even a reasonable difference of 1.5% may produce a location error of up to 1,500 m. This prevents the fault from being located and identified quickly by the service technicians arriving at the location that is equal (route distance) to the (fiber) distance shown by OTDR, as they may be unable to notice that a ditch that has caused the cable to be cut is located approximately 800 m away.

This paper focuses on diagnosing of long (10–130 km) fiber links used in telecom networks, where the accuracy of OTDR-based fault location is of key importance, and where installation techniques and record-keeping facilities used tend to be consistent.

The article is arranged as follows. Section 2 presents the relationships between cable route lengths, cable and fiber lengths, as well as gives the definitions. Section 3 presents a review of fiber optic cables and their helix factors. Section 4 highlights the effects of cable installation on extra lengths of cables and fibers. Section 5 presents OTDR operating principles and issues that are important for the accuracy of fault location, with Section 6 describing how the fiber distance should be converted to cable or route distance in order to facilitate the process of locating faults. Conclusions are presented in Section 7.

2. Optical Fiber Cables – Definitions

This section explains the definitions used: lengths (distances) of optical fiber, cable and cable route, understood as a line joining all facilities between specific terminations of the fiber optic cable link: buildings, manholes, poles, etc.

Several terms used in this paper, such as the "helix factor", "fiber overlength", or "index of refraction", are not fully standardized and are defined differently in the literature, in datasheets, and in OTDR user manuals. To accommodate this, alternative names are indicated throughout the paper.

There are three distinct lengths/distances between a line termination, typically at the optical distribution frame (ODF), where access to fibers for testing purposes is provided, and a cable failure (event) detected by OTDR:

- fiber length: physical length of optical fibers between ODF and the event,
- cable length (sheath length): sum of the length of all cables between ODF and the event,

- route length: cable route length projected onto the ground, as seen on a map, between ODF and the event.

The common rule is:

Fiber length (L_F) > Cable length (L_C) > Route length (L_R).

The difference between fiber and route lengths is usually the largest in aerial networks, where the $L_F : L_R$ ratio may reach 1.08. The other extreme is a cable link with a limited fiber count duct cable of a central loose tube design, and moderate spare segments of cable at each splicing location, say 2×25 m every 2 km, where the $L_F : L_R$ ratio is only 1.006–1.008. Even in this case, the difference between L_F and L_R reaches, after 50 km, 300–440 m.
The relationship shown above is generally applicable due to the following factors:

- outdoor fiber optic cables contain excess lengths of each fiber to accommodate temperature variations and tensile forces without exerting excessive strain on the glass fibers. This is most often done by packing an extra segment into a protective tube a length of fiber slightly longer than this tube, with the fiber forming a helix. Cables with stranded loose tubes or slots include an additional ("dead") length of fibers which cannot be used to compensate for cable elongation;
- joint closures store additional lengths of fibers to allow re-splicing. They are usually 0.5–1 m long on each side of the splice;
- spare lengths of cables, typically 20–40 m, are placed at selected locations along the cable route, preferably near each joint closure to allow repair works;
- aerial lines include vertical runs of the cable at each splicing point;
- duct networks often include numerous road or railroad crossings and offset manhole locations. At each of them, an additional length of duct and cable is provided, e.g. 20 m, but this is not always recorded in network documentation and in the route length data;
- line terminations in big buildings include non-negligible lengths of cables between the entry to the building and ODF ports, typically they are 15–50 m long, while the cable route shown on a map seems to end at the entry to the building.

The above list is not exhaustive due to specific issues, such as wrapping a compact fiber optic cable around a support (messenger wire), which increases cable length.

2.1. Helix and Route Factors

The method used to install the fiber optic cable greatly affects the relationship between the lengths of cable route L_R, cable L_C, and optical fiber L_F. If the network connection follows a uniform set of rules, e.g. consistent lengths of cable sections and spare lengths are used, with one type of cable, the approximate relationship is:

$$L_F = HF \times L_C = HF \times RF \times L_R \ , \tag{1}$$

where HF or "helix factor" (sometimes also referred to as the "cabling factor") is the ratio between the length of fiber in the cable and cable sheath:

$$HF = \frac{L_F}{L_C} \ . \tag{2}$$

HF depends on the type of cable – see Section 3. RF or the route factor is the ratio between the length of cable and the length of route. This parameter accounts for all extra segments of cable in the line introduced by:

- spare lengths of cables,
- vertical sections in aerial installations,
- undulation of directly buried cables to compensate for soil movement,
- running the cable through manholes offset from a straight line, etc.

HF and RF vary within the 1.005–1.04 and 1.01–1.10 ranges, respectively. For example, if a cable section in a duct network is 2,000 m long, and spare lengths of 20 m are stored on each side of each joint closure, $RF = 2040 : 2000 = 1.02$, even with perfectly straight installation ducts. The highest HF and RF values are observed in aerial networks, in particular those with optical ground wire (OPGW) and all dielectric self-supporting (ADSS) cables suspended on high voltage overhead power lines. Such aerial cables are exposed to a wide range of temperatures and tensile loads. In order to accommodate the resulting cable length variations, the cable must have a high overlength of fiber with respect to cable length, with HF of up to 1.04 in cables with stranded loose tubes.
A consistent relationship between route, cable and fiber lengths may be expected in regional or long distance lines built as separate projects. In urban and suburban environments, a much greater variability of RF is observed due to non-straight cable routes resulting from limited rights of way, street layouts, re-use of existing ducts, and obstacles. Hence, only the helix factor is useful in such cases.
While the helix factor is rarely included in cable specifications, it may be calculated from internal dimensions of the cable or may be measured.

3. Fiber Optic Cables

In this section, a review of cable designs is presented to familiarize the reader with differences between the length of fiber in the fiber optic cable (natively indicated by the OTDR) and the cable itself (marked on the sheath), as well as with typical HF values.

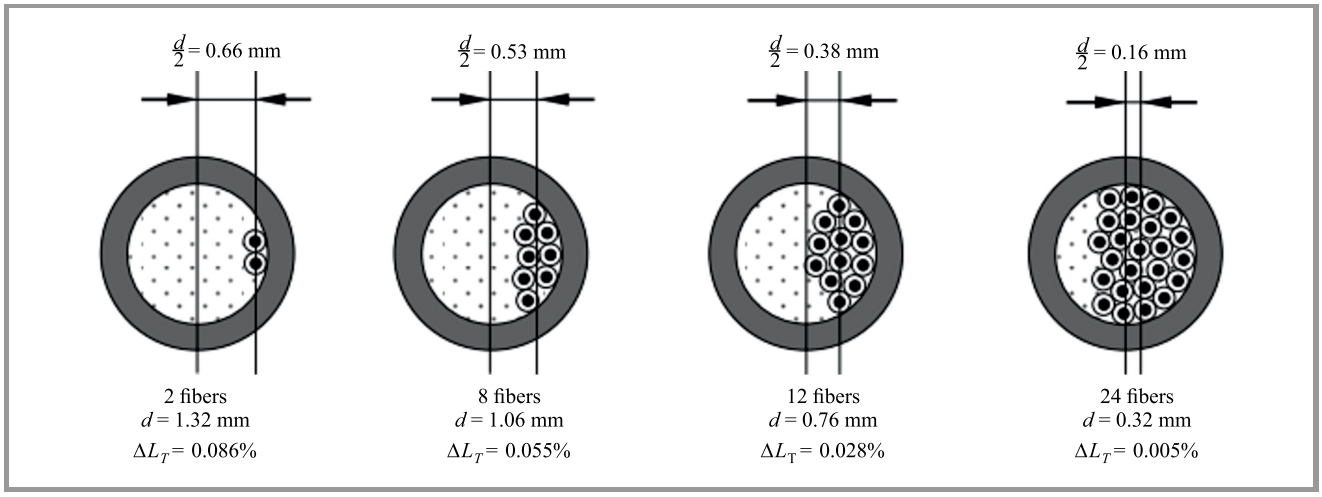


Fig. 1. Approximate helix diameter d and fiber overlength ΔL_T for different numbers of fibers in a 0.25 mm primary coating placed in a 1.6/2.2 mm loose tube. The bundle of fibers is assumed to form a helix with pitch p of 100 mm. All parts are to scale.

3.1. Loose Tube Cables

In outdoor cables, fibers shall be protected against strain, crush, or excessive bending that takes place when the cable length varies with temperature and tensile loads, or when deformation is experienced. A bundle of fibers (colored for identification purposes) [1], [2], is placed in a loosely fitting protective tube, called a loose tube, with a smooth inner surface. To prevent water penetration, the tube is filled with a gel or includes a water-swellable material. Alternatively, the fibers may be formed and glued into ribbons [3] which are placed in cables of either loose tube or slotted core design.

3.1.1. Loose Tube

The glass fibers are mechanically decoupled from the tube by placing an additional, uniformly distributed length of them with respect to the tube. This extra segment is known as fiber overlength or excess fiber length (EFL). It is a result of post-extrusion shrinkage of the polymer tube when it is cooled down from the extrusion temperature of 220–260°C to ambient temperature. Unlike thermoplastic polymers, the fused silica fibers are characterized by a low and almost temperature-independent thermal expansion coefficient within the range of temperatures typical for the manufacturing and use of fiber optic cables, i.e. approx. $0.55 \cdot 10^{-6}$ K for a bare glass fiber and $2.25 \cdot 10^{-6}$ K for a fiber with the standard primary coating with the diameter of 250 μm.

The fibers inside the tube are bent to form a helix, and may be partially or fully straightened when the tube and cable are elongated by tensile force or bent more sharply when the cable contracts in low temperature. As long as the fibers are not fully straightened, they are protected against excessive strain which may cause a failure. However, excessive overlength causes severe bending of fibers and increases attenuation.

The fiber length to tube length ratio is defined by:

$$HF_T = \frac{L_F}{L_T} \ , \tag{3}$$

where: HF_T – helix factor of straight loose tube (≥ 1), L_F – physical length of optical fiber, L_T – physical length of tube.

In some publications and OTDR user manuals, an alternative definition of the helix factor (for tubes and cables) is used, defined as an extra fiber length divided by tube length and expressed as a percentage figure:

$$\Delta L_T = \frac{L_F - L_T}{L_T} \cdot 100\% \ . \tag{4}$$

This parameter is also known as fiber overlength, and this concept will be used here.

If the fiber has a regular helical (spiral) shape with pitch p and diameter d, HF_T may be calculated as:

$$HF_T = \frac{L_F}{L_T} = \sqrt{1 + \left(\frac{\pi d}{p}\right)^2} \ , \tag{5}$$

while fiber overlength in the tube, expressed as percentage rate, is calculated as follows:

$$\Delta L_T = \left[\sqrt{1 + \left(\frac{\pi d}{p}\right)^2} - 1\right] \cdot 100\% \ . \tag{6}$$

For overlengths of up to 1%, encountered in all typical cable designs, a simplified formula is used, retrieved after error correction from [4]:

$$\Delta L_T \approx 493 \cdot \left(\frac{d}{p}\right)^2 \ [\%] \ . \tag{7}$$

Medium and high fiber overlengths, such as 0.1–0.6% ($HF_T = 1.001$–1.006) are desirable in outdoor cables that

are exposed to tensile forces and variable temperatures. The helix diameter d is lower than the inner diameter of the tube, roughly by the width of the bundle of fibers, as shown in Fig. 1. When the tube is filled with fibers, the fibers are positioned straight and overlength is close to zero. This is typical in "tight-fitting" tubes for micro cables or duct cables with very high fiber count.
There is a trade-off between mechanical protection of fibers and their density in the cable. According to the applicable standard [5], a long-term strain of 0.20% is permitted for silica fibers proof-tested at 1% strain. Overlength typical of small diameter loose tubes is not satisfactory for many applications, particularly in aerial networks. Therefore, small tubes are suitable for cables in which stranding results in additional overlength or for cables which are blown into ducts by dedicated pneumatic devices using well-controlled tensile force.

3.1.2. Cable with Central Loose Tube (Unitube)

This type of a loose tube cable incorporates a single, straight tube located in the center and surrounded by sheath, strength members, ripcords, etc. An example is shown in Fig. 2.

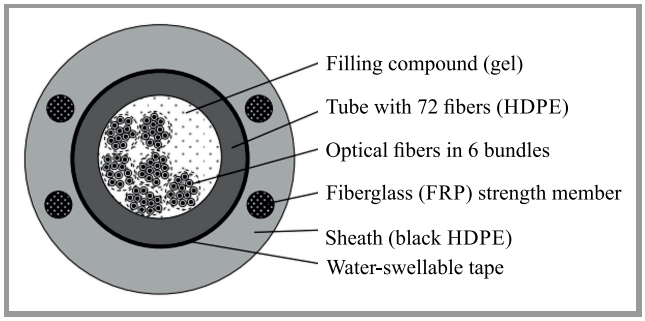


Fig. 2. Cross-section of a duct cable with central a loose tube.

In this case, the overlength of fibers is caused only by forming a helix inside the tube, and:

$$HF = \frac{L_F}{L_C} = HF_T = \frac{L_F}{L_T} \ , \tag{8}$$

where: HF – helix factor of cable, L_C – length of cable, marked on its sheath.

3.1.3. Cable with Stranded Loose Tubes

In this design, the cable core incorporates multiple and identical loose tubes stranded around a central rigid strength member in one or more layers. The number of tubes is 4–12 in a single layer (Fig. 3). Stranding is either helical or reversible, also known as SZ or reverse oscillating lay (ROL). Reversible stranding makes it easier to manufacture cables with the use of more compact machinery and allows to extract selected tubes during cable splicing. The mechanical conditions for fibers are similar. The cable frequently gets a second external strength member made of aramide, glass or basalt fibers to withstand significant tensile forces, especially in aerial applications.

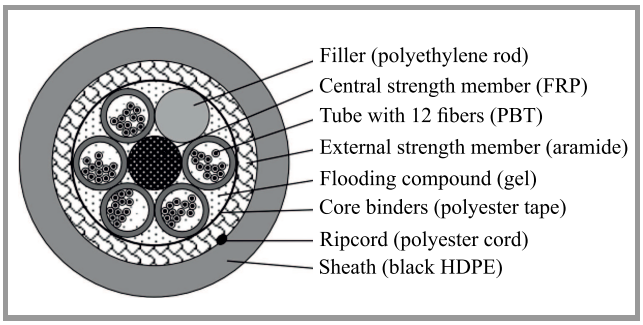


Fig. 3. 60-fiber dielectric, gel-filled duct cable with 5 stranded tubes and filler. All parts are to scale. Fillers enable to strand the core with a reduced fiber and tube count.

Stranding results in overlength of tubes and fibers inside with respect to cable sheath, adding to the overlength of fibers existing in (straight) tubes. For helical stranding, the formulas for the tube helix factor HF_S and overlength are the same as for fibers in the tube:

$$HF_S = \frac{L_T}{L_C} = \sqrt{1 + \left(\frac{\pi d_S}{p_S}\right)^2} \ , \tag{9}$$

$$\Delta L_S \approx 493 \cdot \left(\frac{d_S}{p_S}\right)^2 [\%] \ . \tag{10}$$

Here, p_S and d_S are the pitch and diameter of the helix formed by each tube in the cable. Equations (9) and (10) apply also to helically twisted slots in a slotted core cable. The cable helix factor may be obtained by adding 0.03–0.05% to account for the typical fiber overlength inside a loose tube.
There are several differences with respect to straight loose tube:

- helix diameter d_S is large, typically 4–8 mm in a single layer cable (Fig. 3), and ΔL_S in the range of 0.5-4% may be obtained,
- helix pitch p_S is set during stranding, and ΔL_S is adjusted as desired,
- fibers cannot move all the way to the axis of the cable. Most of the fiber's overlength is "dead", not useful for accommodating cable elongation, but the remaining "net" overlength of 0.15–0.8%, is sufficient.

The resulting helix factor in a cable with stranded tubes is a product of both components:

$$HF = \frac{L_F}{L_C} = HF_T \cdot HF_S \ . \tag{11}$$

For relative fiber overlength, the approximate formula is:

$$\Delta L = \Delta L_T + \Delta L_S \ . \tag{12}$$

The second component is most important, because the tubes have usually a small diameter of 1.5–3 mm.
If the cable contains more than 12 tubes, they are stranded in 2 or 3 layers, e.g. with 6 tubes in the inner layer and 12 tubes in the outer one (Fig. 4), each with a different helix diameter and pitch. The resulting helix factor for fibers in each layer may be different, because the cable is manufactured to achieve an identical "net" overlength and a strain-free window for all fibers. Additionally, manufacturers tend to minimize the length of fibers to reduce costs. A detailed description of the manner in which helix factors are calculated and other related parameters may be found in patents filed by the Corning company [6], [7]. For example, the said description includes the calculation of a complete strain window i.e. cable extension and contraction range with zero fiber strain.

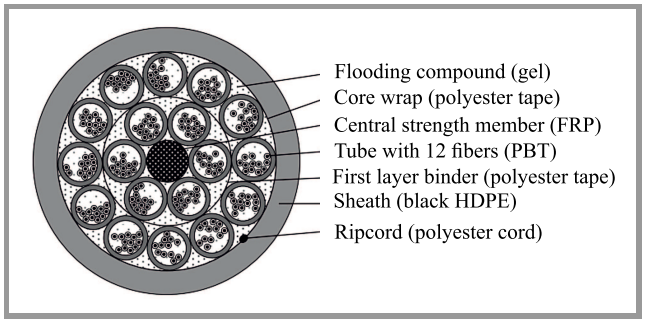


Fig. 4. The cross section of 216-fiber duct cable with 18 (6+12) loose tubes stranded in two layers.

A typical network operation support system (OSS), where data for all cable routes is stored in the network database, usually cannot handle separate HF values for specified ranges of fiber numbers in a single cable, e.g. 1–72 and 73–216 for the cable shown in Fig. 4. Instead, a single HF value for all fibers in a given cable is stored, which causes errors in locating faults for specific parts of the fibers. To overcome this problem, a comment on what range of fiber numbers the HF applies to, e.g. 1–72, and using only OTDR traces of these fibers for locating faults, may be the solution.
Another OSS issue is the need to assign different HF values to each cable section, as future maintenance of the link may be performed with the use of a different cable type.
Duct and directly buried cables need only a moderate net fiber overlength (0.05–0.20%), because they are exposed to large tensile forces only temporarily during installation, and are protected against temperature extremes by thermal inertia of the soil. A compact central tube may be a good approach here.
A different situation is experienced in aerial cables, such as ADSS and OPGW, exposed to variable tensile loads, extreme temperatures and sun heat, where a "net" overlength of 0.4–0.8% is expected. This is realized by adopting a short tube stranding pitch, typically equaling 100–150 mm, as opposed to 300–600 mm in duct cables. Another solution is to use a large diameter central tube, but such a design suffers from other issues, like tube stiffness and migration of fibers inside the tube in aerial installations.

3.1.4. Microcables

A microcable is a compact type duct cable with an outer diameter of 1.2–9.5 mm. While the optical unit is of the loose tube type, major differences may be identified:

- it has only a minimal strength member and a thin sheath,
- it requires blowing with pneumatic machinery and strict control of tensile forces,
- it is installed in plastic microducts of a small diameter (5–14 mm).

Fiber counts reach up to 576 - this is possible due to the use of compact fibers in 200 μm or 180 μm primary coating. The loose tubes in micro cables have small diameters (1.2–2 mm), thin walls and are almost completely filled with fibers. Designs with stranded tubes and a low stranding pitch dominate, but due to the compact optical unit, the HF of a micro cable is usually below 1.015.

3.2. Ribbon Cables

A fiber ribbon is a group of 4–24 fibers in colored primary coatings, laid in parallel and bonded together with a thin, soft matrix (Fig. 5) that may be easily removed mechanically [8]. The ribbon does not include any overlength of fibers.

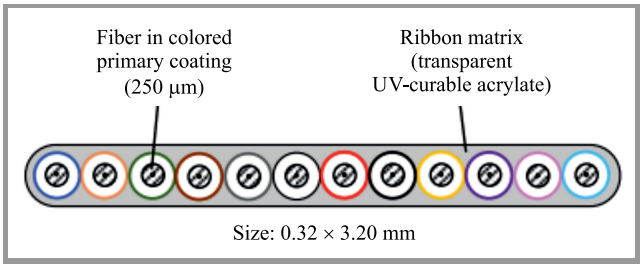


Fig. 5. Cross-section of a fiber ribbon with 12 single-mode fibers.

Back in the day, ribbons were fully coated with a binder material and therefore relatively stiff and difficult to bend or twist. Multiple ribbons of this type may be densely stacked in a large diameter loose tube, or in a rectangular groove in a slotted core cable, with fiber counts of up to 1000 (Fig. 6).
While stacked ribbons fill well rectangular slots, their stiffness restricts undulation that is necessary to obtain overlength of fibers in a tube or slot. A larger net fiber overlength may be obtained in a cable with stranded tubes or a slotted core, where the slots in the central element form a helix. The helix factor of such cables is comparable to that of conventional loose cables with stranded tubes.

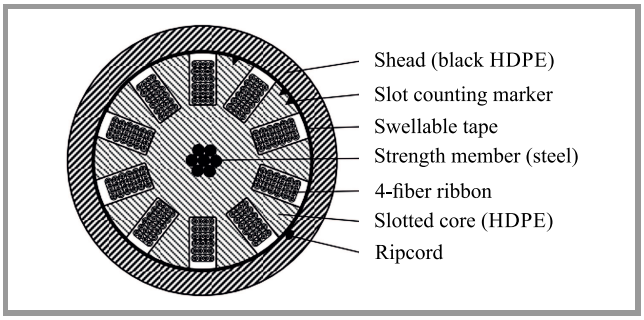


Fig. 6. Schematic of a 200-fiber slotted-core cable with 4-fiber ribbons in rectangular slots.

3.3. High Fiber Count Cables

The recent trend is to develop duct cables with extremely high fiber counts. Such cables have either one or 4–6 large tubes filled with 12-fiber ribbons. The ribbons are of the partially (periodically) bonded type, a solution that allows for easy twisting or bending and ensures that almost 100% of space in the tube is used, also because the "bending tolerant" or "bending insensitive" single mode fibers comply with the ITU-T G.657.A1/A2 standard [9] and, hence, tolerate some mechanical pressure without an increase in attenuation. A good example is a duct cable developed by Fujikura, with the outside diameter of 35 mm and 6,912 fibers in 200 μm primary coating formed into 12-fiber partially bonded ribbons, all in a single central tube.

Similarly to micro cables, this solution is characterized by a low overlength of fibers in tubes (fibers are almost straight). The helix factor may be higher in large diameter cables having stranded tubes.

3.4. Cables with Tight Buffered Fibers

The tight buffer is a layer of a rigid polymer, such as PBT, polycarbonate, polyamide or (moderately) plasticized PVC extruded over a single primary coated fiber – directly or with intermediate layer of a soft polymer, e.g. silicone. The outer diameter of a tight buffered fiber is 0.9 mm (0.036"), and is sometimes reduced to 0.6 mm in compact type cables.

The fiber overlength of 0.05–0.15% (HF = 1.0005–1.0015) is provided by the post-extrusion shrinkage of the outer layer of the polymer, as in a loose tube. The ability to form the fiber into a helix to accommodate overlength is limited by the lack of empty space, and the glass fiber is subjected to compressive strain that decreases with temperature.

The helix factor of an indoor cable with 1–12 straight fibers is higher, e.g. 1.005, because the tight-buffered fibers are surrounded by soft aramid fibers (the strength member) and form a helix or assume a similar shape after the cable jacket has been extruded and cooled. Large capacity indoor cables typically include stranded fiber units, and their HF may be calculated using Eq. (9).

4. Issues with Cable Installation

In this section, we explain how the type of the fiber optic cable and the method of its installation affect the relationship between fiber distance (determined by OTDR) and route distance.

4.1. Duct Installation

This type of outdoor network design is characterized by low requirements in terms of the cables used. The cables are pulled into ducts having the inner diameter of approx. 100 or 37 mm (or even lower in the case of micro cables) and are subjected to moderate tensile loads after installation. Duct cables are protected from temperature extremes by significant thermal inertia of the 80–120 cm thick layer of soil above the ducts, and against crushing forces transferred from the soil. The joint closures and spare segments are stored in underground manholes or handholes. However, the underground environment is frequently humid and cables may be surrounded by water or mud entering the manholes and/or ducts.

Duct cables are designed to withstand moderate tensile loads and temperature ranges. Cables with a single central tube and with multiple stranded tubes, as well as slotted core cables with fiber ribbons (in some countries only) are used here. Their HF ranges from approx. 1.005 for thin unitube cables to 1.015 for large fiber count cables with stranded tubes.

The route factor is the lowest in newly built cable networks located outside of towns, e.g. with cables pulled into plastic ducts laid directly in soft soil. For example, with the length of a cable section of 1500–2500 m, the spare lengths of 20–30 m on each side of the joint closure, with 0.25% of the total length added to account for failing to lay the ducts in an ideally straight line and to take into consideration other factors, the resulting RF equals 1.018–1.042.

In a complex urban environment with multiple obstacles, with existing manholes being used, RF may exceed 1.10. Similar conditions prevail when large spare cable segments, i.e. 200 m and more, are stored in underground containers to allow for splicing work to be performed away from inconvenient terrain. Under such conditions, it makes no sense to apply a uniform route factor. Large spare sections or deviations from a straight route shall be documented as separate objects with precisely recorded lengths and cable length markings.

4.2. Direct Burial

In this type of network design, cables of the strengthened variety, preferably armored, are laid directly in the soil, and therefore are exposed to considerable crush forces and potential soil movements. However, the range of operating temperatures and tensile forces are similar to those experienced in the duct networks, and the design of an optical unit of the cable and its HF are similar.

The route factor is usually low and equals 1.01–1.02, because the spare segments cannot facilitate servicing operations, as moving the cable would require troublesome and costly excavation work. However, the cable is frequently laid in a somewhat undulated form if soil movement, particularly in areas with underground mines, is a known risk.

4.3. Aerial Type of Installation

In the aerial network, each segment of the cable is suspended between two pylons. At both ends of each 2–3.5 km section, the cable enters an enclosure where a spare length L_{SP} = 20–30 m is stored, allowing for splicing work at ground level and for future repairs – see Figs. 7 and 8. In addition, there are vertical runs of the cable at the end of each section, down to the joint closure. For the OPGW cable suspended over the power conductors of a high voltage overhead power line 30–40 m above ground level, a 25 m vertical run is a typical solution. For ADSS cables placed below phase conductors, the vertical runs are shorter. In a high voltage network using OPGW or ADSS cables, the joint closure is usually attached to the pylon 5–15 m above ground to prevent acts of vandalism. When the cables are installed on a medium or low voltage power network or on telecom infrastructure, the vertical runs are short, so we will focus on an OPGW installation on a high voltage overhead power line, considering it to be the worst case scenario.
A section of an aerial cable is presented in Fig. 7.

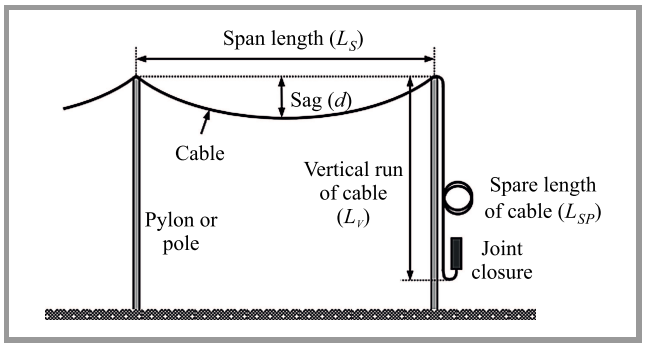


Fig. 7. Schematic of an aerial cable line with key parameters.

The aerial cable forms a catenary line whose length L_A in a section having the span length of L_S and sag (drop) d may be calculated using [10]:

$$L_A = L_S + \frac{8d^2}{3L_S} \ . \tag{13}$$

In this scenario, the cable is subjected to tension. The tensile strain in reference conditions (20°C, no wind, no ice) is approx. 0.15–0.20% for OPGW [11], [12]. To find the true (strain-free) length of the suspended cable in the span, a corrected formula for a reference strain value of 0.2% is suitable:

$$L_A = 0.998\ L_S + \frac{8d^2}{3L_S} \ . \tag{14}$$

Because the fibers in an aerial cable are protected from strain at moderate tensile forces applied to the cable (see subsection 3.1), the sag and its variations depending on temperature, wind and ice hazards have little effect on the physical length of optical fibers in the cable, unless the conditions are extreme.
The typical span in 110, 220, and 400 kVAC overhead power lines in Poland is 300–400 m, and the sag is 3.5–6 m according to [13].
For a 320 m span and a 5 m sag: $L_A = 319.568$ m, and $L_A/L_S = 0.99865$. Counter-intuitively, the actual (strain-free) length of the cable installed in a single span L_A is shorter than L_S, although the difference is low and equals 0.135% only. The extension of the cable exceeds the length added due to the catenary arc formed by the suspended cable. For the following cable and installation specification:

- $HF = 1.035$ (fiber overlength: 3.5%),
- number of spans in a cable section 10,
- route length of a cable section $L_R = 10 \times L_S =$ 3200 m,
- cable spare lengths $2 \times L_{SP} = 2 \times 30$ m (see Fig. 7 and subsection 4.4),
- cable vertical runs $2 \times L_V = 2 \times 25$ m,

we get:

- length of cable in 10 spans: $3200 \times 0.99865 =$ 3195.68 m,
- length of cable in a cable section: 3305.68 m ($RF =$ 1.033),
- length of fiber in a cable section: 3421.38 m,
- $HF \times RF = 1.0692$ (6.92% difference between lengths of fiber and route).

HF in several ADSS cables exceeds 1.04 and the pylons are higher in mountainous terrain, so the difference between measured fiber length L_F and route length L_R may exceed 8%. Even if the location of the fault is referenced to a splice at the beginning of a cable section, it may be difficult to identify span requiring inspection.

4.4. Spare Lengths of Cables and Fibers

These segments allow to perform maintenance without adding a new section of the cable and an extra splice.
The spare length of a duct or aerial cable L_{SP} is preferably stored on each side of the splice and is coiled on brackets in a manhole or on a pole (Fig. 8). If the cable passes an area that is difficult to access with a utility vehicle carrying the equipment necessary for cable splicing, the line may incorporate considerably longer spares, 200 m or more, stored in underground enclosures to allow cable splicing at an accessible location.



Fig. 8. Spare lengths of aerial fiber optic cable accompanying a joint closure box: OPGW on a 400 kV high voltage power line (left) and ADSS on a low voltage line (right).

5. Use of OTDR for Fault Location

An extensive and detailed description of OTDR operation, as well as an interpretation of fiber traces is presented in [14]. E-book [15] is another publication covering this subject, but is considerably less detailed.

The most common kind of fiber cable fault is a localized damage, such as fiber breaks or severe bends, causing high losses in optical fibers. Such failures typically result from excavation and drilling works, as well as from vehicles hitting poles or cabinets. They constitute more than 80% of all fault, usually occur rapidly and affect all fibers in the cable [16].

Fault diagnostics with the use of OTDR include trace acquisition from one end of the affected line to save time. The most common fault, as mentioned earlier, is either a reflective break of the fiber, or a non-reflective and localized loss with uneven values in different fibers. The location of all fiber events is identical and testing multiple fibers brings no improvement in this regards. However, it confirms the type of the fault, e.g. damage of only some of the fibers in the cable suggests a rodent, a shot or a lightning strike.

A specific type of a fiber fault consists in the failure of a contaminated connector or a sharply bent fiber carrying a high-power optical signal in a transmission system employing EDFA amplifiers (up to 0.5 W or +27 dBm) or Raman amplification with powerful (multiple watts) pump radiation injected into the fiber. This kind of fault occurs at the ODF or inside a joint closure located relatively close to a line terminal. It results in the fiber coating overheating and burning out, followed some time later by a break of exposed glass fiber. Both the fiber break and damaged connector are highly reflective events. Other fibers and connectors are not affected.

All OTDRs used for diagnosing telecom networks are capable of automatically analyzing the fiber trace for discontinuities exceeding a set threshold, e.g. 0.10 dB. Such events are automatically listed, along with their type and parameters (loss, reflectivity, fiber distance to event). The fiber trace and the associated event table may be saved to a file or may be uploaded to a remote server. However, the distance displayed is the fiber length L_F between the OTDR's port (or a reference connector, when a launch fiber is employed) and the event. This value must be converted to cable length L_C, using Eq. (2), and next to distance between the fault and the closest easy-to-find objects, such as manholes, cabinets or poles with the joint closure.

The staff tasked with testing the fiber are often subcontractors servicing multiple networks, and do not have any precise data (n_{eff}, HF, route documentation) necessary for precise location of the fault. This information must be provided by the network operator from his OSS or cable network database.

5.1. Instrument Uncertainty in Distance Measurement

The distance uncertainty is usually specified by OTDR manufacturers as a sum of:

- 0.001–0.005% of distance measured (1–5 m for 100 km),
- cursor resolution (0.1–20 m depending on pulse width).

For a distance of 100 km and the pulse width of 1 μs (100 m), the uncertainty defined using this method is below 25 m. With a longer pulse of 10 μs (1000 m), the uncertainty is higher due to the shape of the pulse, optical receiver's impulse response, operator's skill, software used, etc. For a 10 μs pulse we can reasonably expect an uncertainty of approx. 75 m. Still, this is only 0.075% of the fiber distance measured.

The distance measurement error resulting from ignoring the cable helix factor is in the 0.15–4% range, while the spare lengths and vertical runs of cable and undulation of cable may incorporate a comparable error.

5.2. Refractive Index of Optical Fiber

There are two parameters of the optical fiber that are important for fault location using OTDR:

- effective refractive index n_{eff} for calculating fiber length between the OTDR's optical port and the fault,
- attenuation limiting the maximum length of the fiber which may be tested.

The geometry, attenuation, dispersion and mechanical parameters of telecom-grade fused silica fibers are standardized under ITU-T [9], [17] and IEC [18], [19]. The fiber's effective refractive index n_{eff}, defined as the ratio of one-way transmission delay τ in the fiber multiplied by the speed of light in vacuum to the fiber's physical length L_F is not standardized:

$$n_{eff} = \frac{c\tau}{L_F} \ . \tag{15}$$

In general, n_{eff} of multimode fibers (1.47–1.50) is higher than of single-mode fibers (1.46–1.47) because of the lower

difference in the refractive index between the core and the cladding, as well as frequent incorporation of depressed (fluorine-doped) inner cladding in single-mode fibers.
Other names of this parameter include index of refraction (IOR) [20], effective index of refraction (EIOR) [21], effective group index of refraction [22], and group refractive index [23].
Specifications provided by fiber manufacturers include values of n_{eff} at intended operating wavelengths [22], [23] with a decent resolution, e.g. $n_{eff} = 1.4682$. The estimated consistency of n_{eff}, resulting primarily from its dependence on the fiber mode field diameter (MFD) and its production tolerance, typically ± 0.4–0.5 μm, as well as the tolerance of OTDR operating wavelength [21], approx. ± 20 nm, amounts to approx. ± 0.0003. The variations observed with changes in temperature (for fibers in primary coating) and aging are negligible. A margin for other factors like rounding must be added, increasing n_{eff} uncertainty to 0.0001. However:

- n_{eff} is wavelength-dependent due to chromatic dispersion of the fiber, resulting in a difference between values at 1310 nm and 1625 nm of up to 0.002 for the ITU-T G.652 or G.657.A single-mode fiber (for the worst case scenario),
- cable specifications often fail to include n_{eff}, as the fibers in the cable can vary.

The uncertainty in measurement of fiber distance L_F resulting from the uncertainty of n_{eff} is proportional to the measured value. Two scenarios are important here:

a) the user has entered the n_{eff} found in the data sheet of the fiber under test,

b) the n_{eff} value from OTDR factory settings or from previous measurements is used.

The review of available datasheets indicated that n_{eff} values of commercially available single-mode fibers for terrestrial networks, conforming to ITU-T G.652.D, G.654.E, G.655.C/D/E, and G.657.A1/A2 standards, are within the following ranges:

- 1.4606–1.4710 at 1310 nm if the fiber is designed for operation at this wavelength,
- 1.4620–1.4700 at 1550 nm.

The difference between extreme values is about 0.010, so the combined error in setting the n_{eff} in the second case increases to ± 0.0110. Using a n_{eff} value for a wrong wavelength, e.g. 1550 nm instead of 1310 nm, may increase this error up to ± 0.0130.
During the measurement of fiber distance of $L_F = 100$ km, the uncertainty expected in scenarios (a), (b), and (b) with data for wrong wavelength is 68 m (0.068%), 748 m (0.748%), and 884 m (0.884%), respectively. This error in scenario (a) is comparable with distance measurement uncertainty specified by manufacturer of the OTDR.

5.3. Fiber Refractive Index and Cable Helix Factor

For reliable distance measurements, the value of n_{eff} provided by the fiber manufacturer must be set in the OTDR for each wavelength at which fiber testing is planned. The range is at least 1.4–1.6, giving a possibility of introducing a large error by entering wrong n_{eff}, and it may be expected that the value will be reduced to 1.00 due to the recent introduction of low-latency photonic bandgap fibers with a hollow core, designed for data center applications [24]. Fibers with n_{eff} values as low as 1.02 were developed recently as well.
Many OTDRs allows to input cable HF or equivalent fiber overlength as a percentage value [20]. This parameter is wavelength-independent.
Because n_{eff} is usually specified at two or three wavelengths, such as 1310 nm, 1550 nm and sometimes 1625 nm for single-mode fibers, the value at another wavelength, e.g. 1650 nm, may be estimated by linear interpolation or extrapolation [21].
The default value of HF is 1.000 (0.0% fiber overlength). If the OTDR has no HF setting function, a solution is still possible – after setting n_{eff} equal to the n_{eff} of the fiber multiplied by HF, the instrument will show the length of cable instead of fiber.

5.4. Measurement of Cable Helix Factor

The length of an outdoor cable with a rigid strength member, except for the OPGW made of stranded wires, is marked accurately on the sheath by the factory (maximum tolerance of cable length is 0–1% of length markings, while the typical tolerance equals 0.2%). This allows to calculate HF from the fiber length L_F measured with OTDR and cable length L_C between markings at both ends of the length under test, using Eq. (2). However:

- the HF value established using this method is valid only for a given type and size of cable. It varies considerably for otherwise equivalent cables from other suppliers,
- HF is much less consistent than fiber n_{eff}.

6. Distance Correction for Fault Localization

The techniques enabling a more precise localization of fiber cable faults include:

- conversion of the measured fiber length to cable length using the helix factor,
- accounting for spare, vertical and indoor lengths of cables,
- using the route factor which is useful only for long, uniform lines,
- referencing distance to the nearest splice instead of ODF.

Effective use of these methods requires access to OSS information, i.e. a valid database of cable routes and network facilities, including buildings, manholes, poles, street cabinets, splice and termination locations, cables (type, HF value, fiber type, count, and n_{eff}), together with their geographical locations and fiber IDs. Additionally, the OSS shall store the reference OTDR traces of all fibers, measured from both ends of the line, for comparison with the test data acquired before and after servicing.

With a decent set up and well-maintained OSS, the system is capable of calculating the actual location of a fault after entering a raw distance to the fault measured with OTDR, and of presenting this location on a map together with nearby objects, such as manholes or buildings for reference.

Unfortunately, this is not always possible, as the maintenance of infrastructure data is costly and labor-intensive. In such cases, manual or semi-automated correction of distance to a fault may be the only option.

The best way is to set the OTDR up with HF of the cable to be tested. Alternatively, cable length L_C may be calculated from the measured fiber length L_F using Eq. (2).

This method has no use when the line is made up of two or more different types of cable, unless corrections for each part are made separately, which is complex and with inherent error risk.

The use of route factor (RF) is recommended only for long, uniform lines with one type of cable, uniform installation rules, straight routes, and even so, specific "jitter" resulting from the conversion of discrete spare and vertical lengths of cables to a fraction of total cable length is experienced, meaning that the accuracy of this method is low. The length of route to fault L_R may be calculated from measured fiber length L_F using Eq. (1).

Referencing to the nearest splice is the best approach when combined with use of cable helix factor for conversion of fiber distance to cable distance. While a complete single mode fiber optic link may have a length of up to approx. 120 km, the cable section extending between splices is typically up to 4 km long, and most often 2.5 km or even less. Consequently, errors in the measurement of distance to fault stemming from an unknown or improperly set HF and n_{eff} are proportionally reduced.

The suggested fault location procedure is as follows:

- acquire a fiber trace and set the cursor at the beginning of a fault (spike or fall on the fiber trace),
- look for the nearest splice before the fault and measure the distance between them,
- verify whether this distance is smaller than the length of the cable section. If not, this means that the nearest splice was missed due to very low (apparent) loss, and another fiber shall be tested,
- convert the measured splice-to-fault fiber distance to cable distance, using Eq. (2),
- use this cable distance to find the relative location of the fault with respect to the nearest objects in the route documentation, e.g. the fourth manhole after the one with a joint closure. The spare lengths of cable must be included.

In a duct network with a 2000 m cable section, cable with $HF = 1.015$ (1.5% fiber overlength), 0.2% uncertainty of HF (0.002), and spare length of cable near the joint closure $L_{SP} = 20$ m, the expected error in a fault location process relying on this procedure is approx. 5 m. Without corrections for HF and spare lengths of cable, the error may reach 50 m under the same conditions.

7. Conclusions

Precise location of a fault in an outdoor fiber optic cable using an OTDR is crucial for enabling fast and efficient cable repairs and for restoring the fiber connections. The use of OTDR allows to locate a fault from a distance of 100 km or more, with the accuracy in order of 100 m, therefore enabling to begin servicing without losing any time on looking for the actual fault location.

The methods of estimating the locations of fiber cable faults, as presented in this paper, shall be useful in achieving this goal, especially when a fully featured network OSS is not implemented. The use of the parameter of fiber optic cables discussed in this paper, i.e. the helix factor, is always essential.

References

[1] IEC 60304 Ed. 3,"Standard colours for insulation for low-frequency cables and wires", 19821 [Online]. Available: https://webstore.ansi.org/Standards/IEC/IEC60304Ed1982

[2] ANSI/TIA-598-D-2014, "Optical Fiber Cable Color Coding", *Telecommun. Industry Association (TIA)*, 2018 [Online]. Available: https://global.ihs.com/doc_detail.cfm?document_name=TIA%2D598&item_s_key=00134525

[3] IEC 60794-1-31 Ed 2, "Optical fibre cables – Part 1-31: Generic specification – Optical cable elements – Optical fibre ribbon", 2021 [Online]. Available: https://webstore.iec.ch/publication/30731

[4] H. Murata, *Handbook of Optical Fibers and Cables*, 2nd Ed. New York: Marcel Dekker, 1996 (ISBN: 9780585268972).

[5] IEC 60794-1-20, "Optical fibre cables – Part 1-20: Generic specification – Basic optical cable test procedures – General and Definitions [Online]. Available: https://webstore.iec.ch/publication/3479

[6] Patent WO 03/083518 A2, "Optical fiber cable with controlled helix-plus-EFL values and methods therefor", 2003 [Online]. Available: https://patents.google.com/patent/WO2003083518A2

[7] Patent US 6859592 B2, "Optical fiber cable with controlled helix values", 2005 [Online]. Available: https://patentimages.storage.googleapis.com/ce/c8/a1/926b52d292150c/US6859592.pdf

[8] IEC 60794-1-31 Ed 2, "Optical fibre cables – Part 1-31: Generic specification – Optical cable elements – Optical fibre ribbon", 2021 [Online]. Available: https://webstore.iec.ch/publication/30731

[9] ITU-T G.657, "Transmission media and optical systems characteristics – Optical fibre cables – Characteristics of a bending-loss insensitive single-mode optical fibre and cable", 2016 [Online]. Available: https://www.itu.int/rec/dologin_pub.asp?lang=s&id=T-REC-G.657-201611-I!!PDF-E&type=items

[10] ITU-T L.89, "Design of suspension wires, telecommunication poles and guy-lines for optical access networks", 2012 [Online]. Available: https://www.itu.int/rec/T-REC-L.89-201202-I

[11] Installation Instructions – AFL Optical Ground Wire (OPGW). AFL, ACS-WI-805.06, 2012. [Online]. Available: https://www.bidnet.com/bneattachments?/689129089.pdf

[12] S. Karabay, E. A. Güven, and A. T. Ertürk, "Performance testing of an optical ground wire composite", *Materials and Technol.*, vol. 47, no. 1, pp. 119–124, 2013 [Online]. Available: http://mit.imt.si/izvodi/mit131/karabay.pdf

[13] D. Złotecka, "Nadmierne zwisy linii WN jako przyczyna awarii w systemach elektroenergetycznych", *Przegląd Elektrotechniczny*, vol. 94, no. 10, pp. 143–147, 2018 [Online]. Available: http://pe.org.pl/articles/2018/10/33.pdf [in Polish].

[14] D. R. Anderson, L. Johnson, and F. G. Bell, "Troubleshooting optical-fiber networks: understanding and using your optical time-domain reflectometer", 2nd Ed. *Elsevier Academic Press*, 2004 (DOI: 10.1016/B978-0-12-058661-5.X5020-4).

[15] J. Laferriere, G. Lietaert, R. Taws, and S. Wolszczak, "Reference guide to fiber optic testing", *Viavi Solutions*, vol. 1, 2018 [Online]. Available: https://www.normann-engineering.com/support/white_papers/fiberguide1-bk-fop-tm-ae.pdf

[16] R. Dutta and G. Rouskas, *CSC/ECE 791B – Survivable Networks Introduction*, North Carolina State University, Raleigh, N.C., USA: Spring, 2008.

[17] ITU-T G.652, "Characteristics of a single-mode optical fibre and cable", 2016 [Online]. Available: https://www.itu.int/rec/T-REC-G.652/en

[18] IEC 60793-2-50, "Optical fibres – Part 2-50: Product specifications – Sectional specification for class B single-mode fibres", 2015 [Online]. Available: https://webstore.iec.ch/preview/info_iec60793-2-50%7Bed5.0%7Den.pdf

[19] IEC 60793-2-10, "Optical fibres – Part 2-10: Product specifications – Sectional specification for category A1 multimode fibres", 2019 [Online]. Available: https://webstore.iec.ch/publication/62020

[20] OTDR – Optical Time Domain Reflectometer User Guide, EXFO Inc., 2019 [Online]. Available: https://trsrentelcoimagerepository.azureedge.net/trswebsitedata/Specs-Manuals/max740c%20manual-10779-11014.pdf

[21] "Explanation of the sources of variation in optical fiber effective group index of refraction values", *Corning Inc.*, AN4091, 2020 [Online]. Available: https://www.corning.com/content/dam/corning/media/worldwide/coc/documents/Fiber/application-notes/AN4091.pdf

[22] Corning SMF-28 Ultra Optical Fiber Product Information. PI1424, 2014 [Online]. Available: https://www.corning.com/media/worldwide/coc/documents/Fiber/SMF-28%20Ultra.pdf

[23] "AllWave optical Fiber – Zero Water Peak: The industry's first zero water peak single-mode fiber for reliable full-spectrum performance (data sheet)", *OFS Fitel*, ID: fiber-117, 2017 [Online]. Available: https://fiber-optic-catalog.ofsoptics.com/documents/pdf/AllWave+-ZWP-Fiber-159-web.pdf

[24] "AccuCore HCF Fiber Optic Cable and Assemblies (data sheet)", *OFS Fitel*, 2020 [Online]. Available: https://www.ofsoptics.com/wp-content/uploads/AccuCore-HCF-Fiber-Optic-Cable-Assembly-web.pdf



Krzysztof Borzycki received his M.Sc. in Electrical Engineering from Warsaw University of Technology, Warsaw, Poland in 1982, and Ph.D. degree in Communications Engineering from the National Institute of Telecommunications (NIT), Warsaw, Poland in 2006. He has been with NIT since 1982, except for the time spent on developing DWDM solutions at the Ericsson AB R&D Center in Stockholm, Sweden, in 2001–2002. He is currently an Assistant Professor at the NIT Central Chamber for Telecommunication Metrology. His areas of interest include fiber access networks (FTTx), testing and standardization of fiber cables and passive components, monitoring of fiber and copper cable networks, security of optical networks and effects of high temperatures in fused silica fibers. Dr. Borzycki has participated in multiple European research programs, including COST-270, COST-299, COST TD1001 and NEMO, and is a member of Cables and Fiber Optics Work Groups of the Polish National Standardization Committee (PKN).

https://orcid.org/0000-0001-6066-6590
E-mail: k.borzycki@il-pib.pl
National Institute of Telecommunications
Szachowa 1
04-894 Warsaw, Poland



Paweł Gajewski received his M.Sc. degree in Fine Mechanics from Warsaw University of Technology, Warsaw, Poland in 1983, and has been an employee of the National Institute of Telecommunications, Warsaw, since. His professional activities focus primarily on designing and programming of measurement, monitoring and control equipment and systems. He has always been focusing on hardware design and development of software for specialized microprocessor systems, in particular for measurement and control applications in various fields of technology. Mr. Gajewski was involved, inter alia, in the development of a system for monitoring the operation of ASM-type telephone exchanges, in SMOK, SPOT, and SMIT development projects focused on monitoring telecom and industrial cable networks, in the development of remotely controlled fiber optic switches for cable monitoring, as well as in the design of equipment for measuring the quality indicators of AC power systems. He is an author of 5 patents and author or co-author of 30 publications.

https://orcid.org/0000-0003-0931-2476
E-mail: p.gajewski@il-pib.pl
National Institute of Telecommunications
Szachowa 1
04-894 Warsaw, Poland

Paper

Fusion of Depth and Thermal Imaging for People Detection

Weronika Gutfeter and Andrzej Pacut

Biometric and Machine Intelligence Laboratory, Research and Academic Computer Network (NASK)

https://doi.org/10.26636/jtit.2021.155521

Abstract—The methodology presented in this paper covers the topic of automatic detection of humans based on two types of images that do not rely on the visible light spectrum, namely on thermal and depth images. Various scenarios are considered with the use of deep neural networks being extensions of Faster R-CNN models. Apart from detecting people, independently, with the use of depth and thermal images, we proposed two data fusion methods. The first approach is the early fusion method with a 2-channel compound input. As it turned out, its performance surpassed that of all other methods tested. However, this approach requires that the model be trained on a dataset containing both types of spatially and temporally synchronized imaging sources. If such a training environment cannot be setup or if the specified dataset is not sufficiently large, we recommend the late fusion scenario, i.e. the other approach explored in this paper. Late fusion models can be trained with single-source data. We introduce the dual-NMS method for fusing the depth and thermal imaging approaches, as its results are better than those achieved by the common NMS.

Keywords—depth imaging, person detection, sensors fusion, thermal imaging.

1. Introduction

The primary goal of this work is to explore the feasibility of detecting human silhouettes in non-visible light spectrum images, without accessing RGB images for reference. Our experiments focus on thermal and depth images showing people in indoor and outdoor environments, i.e. images that are similar to surveillance footage.

The use of alternative imaging sources in computer vision-related tasks is important for numerous reasons. One of those reasons is that they extend the spectrum of features that can be recognized. In this case, recognition is based on the temperature of objects (thermal imaging) and on their geometrical features (depth imaging). Temperature measurements may be critical in the context of the recent pandemic and the demand for wide-scale systems capable of monitoring health parameters. Privacy is another essential aspect that needs to be taken into consideration. At certain location, the use of standard RGB cameras may be prohibited to protect the privacy of data subjects. In such circumstances, surveillance systems relying on alternative vision cameras may prove to be the best solution available. When working with alternative vision systems, the fact that fewer resources are available than in the case of RGB-image based architectures (data needed to train the algorithms or evaluation benchmarks) is the key challenge.

Consumer-grade non-RGB detectors are usually characterized by lower resolution levels, and the images have poorer quality than their RGB counterparts. That may affect the precision of detection. The size of thermal images used in this paper is 160×120 pixels and the average size of the detection boxes framing human silhouettes is approximately 36×56 pixels. Still, detection performance of thermal images surpasses that of higher resolution depth images (see Tab. 1) with the resolution of the latter equaling 1280×720 pixels. This is probably caused by a higher level of noise in depth images which hampers their depth accuracy. Therefore, a fusion of different image sources may lead to the improvement in results. Specific methods relied upon for merging thermal and depth imaging will be discussed in the second part of the paper.

The investigation of the ability to determine the correct fusion methodology required that a dataset be identified containing images of both types, with the pairs of images being spatially and temporarily aligned. These requirements are met by the IPHD dataset which was complied for the Identity Preserving Human Detection Challenge [1] organized in 2020. The IPHD dataset was built using frames extracted from two synchronized image streams: a thermal one and the other containing depth-related information. The dataset is used to evaluate the detection methods proposed in this work and to train the models. We do not use any auxiliary data for estimating models' weights. However, we employ transfer learning techniques from models that were pre-trained using the Common Objects in Context dataset [2].

2. Dataset with Thermal and Depth Images

The IPHD [1] dataset was compiled by researchers from the Chalearn Looking at People group. The entire set consists of over 100,000 pairs of images cut from two video streams, without maintaining information about their order. The footage was captured indoors and outdoors, at such places

Table 1
Detection evaluation metrics for single depth and thermal models based on the Faster R-CNN architecture computed based on the IPHD-test dataset. For reference, results of a method proposed by the authors of the IPHD database are shown as well. As no MAP was used in their publication, the relevant fields are marked NA (not available)

Method	Detection precision (± std. dev.)					
	Thermal images			Depth images		
	AP_{50}	AP_{75}	MAP	AP_{50}	AP_{75}	MAP
Single model (ours)	83.23% (±0.28)	56.38% (±0.32)	51.77% (±0.29)	71.26% (±0.59)	44.01% (±0.45)	42.46% (±0.49)
Baseline method from [1]	52.45%	15.95%	NA	34.32%	9.91%	NA

as streets, universities, libraries, and private houses. The video frames were spatially and temporally aligned. Therefore, they may be used independently or may be merged using one of the fusion methods. The extracted images were arbitrarily split into the development set (IPHD-dev), consisting of 84,818 images of both types, the validation set (IPHD-val), consisting of 12,974 images of both types, and the test set (IPHD-test), containing 15,115 images.

All the images are associated with ground-truth bounding boxes that show the position of people in the scenes. People visible in the images perform various actions: sitting on the sofa, lying on the floor, cooking, eating, talking on the phone. The scenes were manually labeled by the authors using RGB images. The RGB stream was also aligned with the other two, but it is not a part of the publicly available set. Manual labeling would be much more difficult or even impossible for depth or thermal images, since human body features are not easily distinguishable "with the naked eye". One may notice (see Fig. 1) that the temporal alignment is imperfect, and some of the ground-truth detection boxes should be slightly shifted, especially the ones with individuals moving fast. This is caused by issues with the synchronization of sensors at a hardware level. Images in the test set (IPHD-test) were manually adjusted to compensate for the misalignment. The effect of label weakness will also be discussed in this work. Because of the inferior human perception of alternative imaging sources, they are often considered privacy-preserving. However, the extent to which they do not contain any individual features needs to be investigated further.

2.1. Thermal Imaging

In thermal imaging, individual pixels represent far-infrared radiation measured by an IR detector. Radiation may be either emitted or reflected from the scene. The type of representation and its range vary across different sensors, but most thermal detectors have significantly lower resolution than DSC and video cameras used in other imaging systems. Images that are included in the IPHD dataset were collected with the FLIR Lepton 3 sensor. It is capable of detecting infrared waves with the length of 8 to 14 μm [3]. The original resolution of the sensor is 160×120 pixels. The thermal images in IPHD were padded to ensure that their ratio is consistent with that of depth images. As a result, they are 1-channel 16-bit pictures with a resolu-

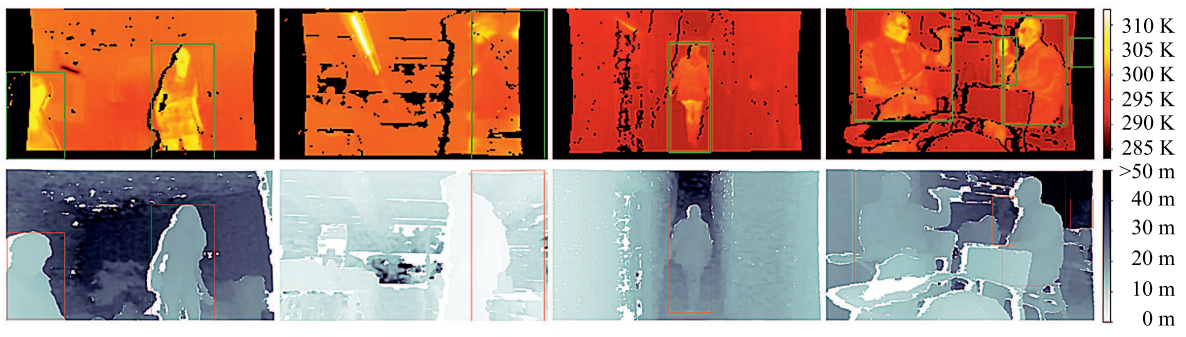


Fig. 1. Spatially and temporally aligned frames from the IPHD dataset: thermal images (upper row, in Kelvin) and depth images (bottom row, in meters). The images are shown in the altered color scales for better visualization (originally, one-channel 16-bit images). The bounding boxes indicate ground-truth locations of people (green boxes for thermal images, red boxes for depth images). Ground-truth labeling of non-RGB images can be challenging, because some of the features are difficult to perceive by the human eye. Here, in the IPHD dataset, labeling was performed using the corresponding RGB stream, which was not publicly available. Nevertheless, a temporal misalignment may be noticed in some of the bounding boxes (second column). The misalignment is caused by problems with synchronization between RGB and other sources. In the test part of the dataset (IPHD-test), the positions of boxes were adjusted manually. Therefore, this subset contains stronger labels than those in the training part (IPHD-dev). (see the digital edition for color images)

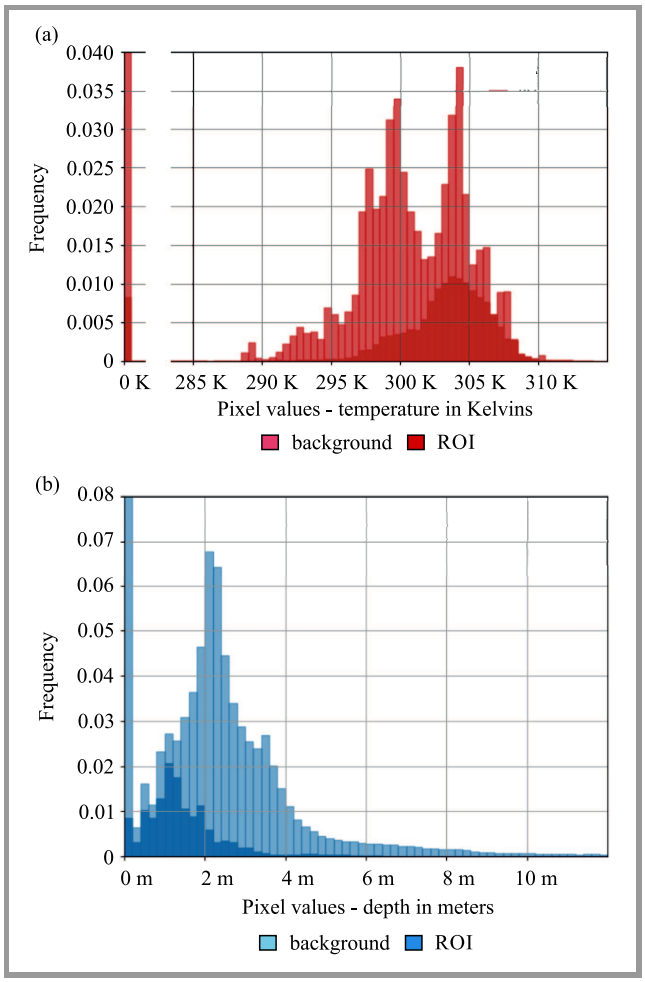


Fig. 2. Histograms of pixels from the IPHD dataset, computed separately for the areas occupied by human silhouettes (ROI) and for the background pixels. The temperature of pixels in thermal images (a) may vary from 0 to 450 K, with most values being within the 290–310 K range. Some of the pixels are set to 0 because of the spatial registration with depth images. Null pixels also denote depth data missing. The values of pixels in the depth images (b) range from 0 to 65 m, where 0 is typically caused by erroneous readings. Most people are positioned not further than 4 m from the camera.

tion of 213×120 pixels. Each pixel represents a specific level of temperature in the Kelvin scale. The padded portions of the images are filled with zeros, and zeros appear also at locations where the depth images render erroneous readings.

Temperature distribution of the images is presented in Fig. 2a. Histograms were computed independently for the areas occupied by human silhouettes and for the background. It may be observed that most of the pixels representing humans have temperatures in the range of 295 K to 310 K, with a mean value of approx. 303 K.

2.2. Depth Imaging

Depth cameras have become more popular and available due to the abundance of devices serving as game controllers. Depth sensors may be easily used to map the pose of a human silhouette and other parts of the body, e.g. hands. Although depth imaging may be based on various hardware architectures, a typical consumer depth camera, like Microsoft Kinect v1 or Intel RealSense, consists of an active infrared projector and at least one infrared detector. The projector casts an invisible light pattern onto the scene. The distance of the object from the detector is estimated using the triangulation method that measures light pattern displacements [4].

The IPHD dataset contains depth images acquired with the use of the Intel RealSense D435 sensor. The size of the images equals 1280×720 pixels [5]. The images are 16-bit channels, just like their thermal counterparts. Pixel values represent the distance to the depth sensor, expressed in millimeters and have the maximum value of 65 meters. Null pixel values express erroneous readings. Bad pixels may appear at the borders of objects or human silhouettes and may also be caused by reflective materials or strong illumination.

Histograms of pixel values calculated for unprocessed depth images are shown in Fig. 2b. As far as the temperature is concerned, they were made separately for the background pixels and the pixels assigned to the human body. It may be observed that there are no labeled human subjects in the range greater than 4 meters: at these distances, the bounding boxes would be too small, and their contents would not be easily distinguishable.

2.3. Image Preprocessing

Before deploying the detection algorithms, the datasets were examined to choose the best preprocessing method. For the preliminary observations, a small subset of 50 images was drawn from IPHD-dev. It is hereinafter referred to as IPHD-pre. Using IPHD-pre, we manually labeled the masks that indicated precise people locations. The masks were used to select two distinct parts of the image: the foreground in which a person appears (ROI) and the background. The histograms of pixel values for those two subareas are shown in Figs 2a and 2b, respectively. As the IPHD database authors suggested, the pixel values in thermal images should be standardized before further processing. In our work, the pixels in thermal images were clipped at the minimum value of $x_{min} = 285$ K and the maximum value of $x_{max} = 315$ K. Then, the images were normalized using the mean and standard deviation calculated on the full set of ROI pixels, excluding the null-value pixels. More precisely, each thermal image was preprocessed by extracting mean value $\tilde{x}_{th} = 296.4$ K and dividing it by $\sigma_{th} = 330$ K such as:

$$X[X > x_{max}] = x_{max} \ ,$$
$$X[X < x_{min}] = x_{min} \ ,$$
$$X = \frac{X - \tilde{x}_{th}}{\sigma_{th}} \ .$$

The depth images were also standardized before processing them with the use of the detection algorithms. Namely, they were clipped to match the range of $[0,\ldots,12\text{ m}]$ and normalized with $\tilde{x}_d = 2.830$ m and $\sigma_d = 3.198$ m.

3. Human Detection Methods Applied to Depth and Thermal Images

There are several reasons why to include either depth or thermal imaging sources in the detection system. One is to augment the information about the objects for their further classification. The second is to leverage the precision of detection by adding potentially complementary imaging features. Finally, the third reason is to diminish the influence of bad lighting or image noise. The methods proposed in the literature are based on various combinations of image types: RGB and thermal [6], [7], RGB and depth [8], [9], [10] or, less commonly, thermal and depth [11].

This kind of a system is found to be functional in many real-world applications, from vision for autonomous driving [12], to industrial inspection [13], monitoring of car and vessel traffic [14], [15], drone surveillance [16], to pedestrian detection. In our work, we focus on the task of automatic person detection. Human detection solutions are widely discussed in the literature, but mostly in the context of RGB imaging. Most of the RGB-based methods apply also to alternative imaging sources. Older approaches are based on local descriptors, such as HOG or SIFT features [17], [18], which have to be hand-crafted. The recently proposed methods use predominantly convolutional neural networks inside their detection pipeline [19]. The most popular human detection architectures detection are Faster-RCNN [20], Mask-RCNN [21], SSD [22] or YOLO [23].

The authors of the IPHD database introduced the detection model based on YOLO in their paper [1]. A YOLO network is a single-stage detector that is available in various implementations, with its third version described in paper [23] still being one of the most commonly used variations. The IPHD baseline method for thermal-depth detection proposed a middle fusion network in which images are merged at the second-to-last convolutional layer level. The results generated by the algorithm serve as a point of reference for comparison with other methods, as they were computed under identical conditions and on the same dataset as in the case of our experiments.

The method proposed in our work is an extension of the Faster R-CNN method introduced in paper [20]. We have chosen this detection network to be the core of our system because it is easy to modify and perform better in terms of accuracy than YOLO (based on the results presented in [23]), as we do not intend to take into consideration other metrics, such as performance. Faster R-CNN involves a 2-stage detection procedure consisting of the region proposal network (RPN) and the second stage module responsible for final object detection and classification. RPN produces a set of proposals that are then fine-grained. In the previous solution, named Fast-RCNN, these two stages were implemented by two different networks, but in Faster-RCNN, all functions are realized by one module. In our implementation, the Faster R-CNN network is built using the ResNet-50 [24] module (Fig. 3). The loss function has two components. One is the classification loss L_{cls} for assigning the probability of the object belonging to one of the classes. Here, as we perform solely person detection, the estimation distinguishes two classes only. The other is regression loss L_{reg} that compares the coordinates of box t_i with the ground-truth coordinates t_i^* where i is the box index, p_i is the measure of "objectness" and p_i^* is equal to 0 in the event of false detection.

$$L(\{p_i\},\{t_i\}) = \frac{1}{N_{cls}}\sum_i L_{cls}(p_i,p_i^*) + \lambda\frac{1}{N_{reg}}\sum_i p_i * L_{reg}(t_i,t_i^*). \tag{1}$$

Typically, at the end of the detection procedure in R-CNN networks, some post-processing methods need to be applied to reduce the number of regions that overlap too closely. Non-maximum suppression (NMS) is one of the most commonly used algorithms, and its will be elaborated on in the section concerned with late fusion.

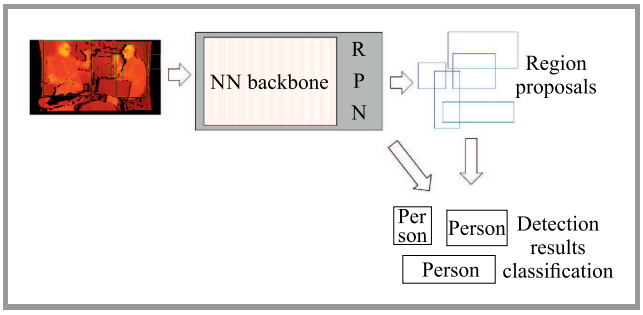


Fig. 3. Single detection model based on Faster-RCNN. Faster-RCNN typically consists of the RPN that generates a candidate list of detection boxes. RPN may be realized by one of the selected convolutional neural network architectures, such as ResNet-50 or AlexNet. At the end of data processing, ROI pooling is performed on the list of candidates to produce the final list of results with classification scores.

4. Results and Discussion

Average precision (AP) is typically used as the evaluation metric to test the detection methods. AP may be computed at different ranges of overlap between detection results and ground-truth bounding boxes. The level of precision with which two sets of coordinates are capable of describing the same object is measured by the intersection-over-union (IOU), defined by the area of overlap between two bounding boxes divided by the area of union.

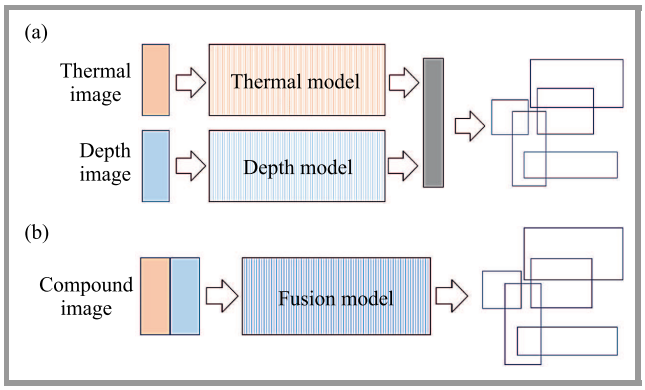


Fig. 4. We experimented with different model fusion patterns, namely late fusion (a), where the results of two models are merged into one set of results using dual non-maximum suppression, and early fusion (b) – where the images from thermal and depth streams are combined into one input of the model trained to recognize the compound input.

The authors of the IPHD dataset suggested AP_{50} (average precision at IOU=50%) to be the primary evaluation metric of the challenge. This choice was motivated by the weakness of the ground-truth labels. Typically, in detection benchmarks, AP_{75} is used, and it was calculated in our experiments as well (average precision at IOU = 75%). We also added MAP (mean average precision) as defined in the COCO challenge [2], being the mean value of AP at IOU=[50%: 5%: 95%].

Our experiments started by training two independent models based on Faster-RCNN networks, with ResNet-50 serving as the backbone (Fig. 4). The first one was trained on thermal images. Random crops and horizontal flips augmented the training set. Network weights were optimized with the use of the SGD method, with the learning rate initialized at 0.005 and updated every three epochs. The thermal detector achieved AP_{50} equal to 83.23% and AP_{75} equal to 56.38% (see Table 1). All the results are computed on the test part of the IPHD dataset (IPHD-test) and are averaged for three repetitions of the training and evaluation procedures.

The second model – with its architecture identical to the first one - was trained on depth images. For this setup, we obtained $AP_{50} = 71.26\%$ and $AP_{75} = 44.01\%$, respectively. Detection evaluation metrics for the depth model were much lower than those for the thermal network. Therefore, we may consider the depth data to be more challenging than their thermal counterparts for running the detection algorithms on.

Some qualitative results of experiments with single-model detection are shown in Fig. 5. We can see that for the depth model, there is often a more significant number of false detections (third column in Fig. 5) and missed detection, especially for smaller objects (second column in Fig. 5). Detection errors may be easily explained, since some objects are difficult to distinguish from people in the depth scenes for the human observer. On the other hand, we noticed false detections pointing to a dog visible in the thermal images. It would be probably reduced by adding labels of other warm objects (like animals and electrical appliances) to teach the network to distinguish them from humans.

4.1. Fusion of Models

Fusion of deep neural models is a technique that can typically boost the accuracy of results, as shown in paper [25] in relation to the classification of videos. The authors distinguished there three approaches to classify the content of multi-frame data: early fusion, late fusion, and slow fu-

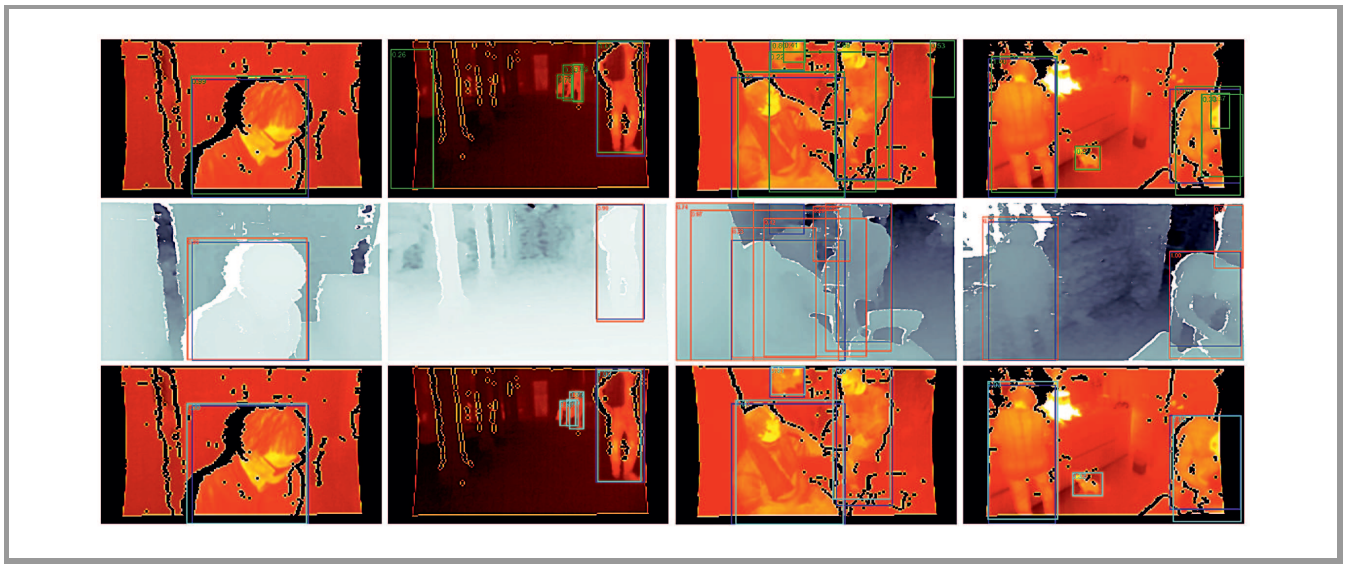


Fig. 5. Qualitative human detection results using IPHD-test data. Blue rectangles indicate ground truth bounding boxes. The consecutive rows show the detections for a single thermal model (first row, green boxes), the detections for a single depth model (second row, red boxes), and predictions from the model with early fusion (third row, cyan boxes).

sion. The division is defined by the point of the network at which the flow of information is merged. Early fusion means concatenating the data before further processing it inside a model. Consequently, in late fusion, the data is merged at the end of the model. In slow fusion (or middle fusion), merging is performed at some of the intermediate layers of the model. The slow fusion technique was applied to the detection network in the IPHD method discussed here for reference (results showed in Tab. 1). Also, the two other approaches (late and early fusion) can be transferred directly to the domain of multimodal images. Therefore, we decided to employ them in our experiments. Late fusion was realized by the application of different versions of the non-maximum suppression algorithm (NMS). Models trained in the previous experiments were reused as components of the fusion system. In the case of early fusion, it was necessary to change the structure of the model and perform the entire training procedure. Results for the fusion strategies were compared against results for the single image approaches discussed in the preceding section.
The code of the library was written with PyTorch and it is available online: `github.com/weronikagutfeter/Red-Hot-Deep-Blue`.

4.2. Dual-NMS

Our late fusion method was based on non-maximum suppression (NMS) [26]. NMS is a post-processing algorithm that is typically used to minimize the number of redundant and overlapping detection results. Simple non-maximum suppression, also called greedy NMS, begins with sorting the detection boxes by their scores in the descending order. Then, results from the sorted list are compared, one by one, with the remaining results. If the IOU of the compared pair of boxes is larger than a selected threshold value, the box with the lower confidence score is removed.
This paper employs a modified version of the NMS algorithm to merge the detection boxes originating from the two distinct models: thermal and depth. The modified version of the algorithm is called dual-NMS. The idea behind this approach is to collect pairs from the two lists of detection boxes, which are also sorted by their confidence scores, like in the simple NMS method. Boxes with the highest scores (from the depth or the thermal lists) are taken, one by one, and compared with all boxes from the other list. The selected detection box is paired with the result with a sufficient IOU and the highest score among the candidates from the other list. The pair is then merged into a single result, and the final detection box coordinates are updated by applying weighted averaging of the coordinates of the components.

Since some unassigned detection boxes may be left after the pairing, several approaches to managing unpaired boxes were evaluated. The simplest solution is to remove the unassigned results, as they are either not present in any stream, have low confidence scores, or are duplicates. However, one of the streams is likely to be a more robust source of detection results. Thus, we may leave the unpaired samples from this source.

To sum up, we checked four versions of the algorithm: with all of the unpaired results removed, with all of the unpaired results kept, with only the unpaired thermal results kept, and, finally, with the unpaired depth results kept. For the reference, we compared dual-NMS with the simple NMS algorithm applied to the concatenated list of detection boxes. Precision rates of the fusion are presented in Table 2. For comparison, the results for single model detectors from Table 1 are shown at the bottom. The best results for late fusion approaches were obtained for the dual-NMS with thermal results kept, for which $AP_{50} = 83.31\%$ and $AP_{75} = 57.84\%$. However, this model was only slightly better than the single thermal model. It can be concluded that for the NMS-based system, the impact of the depth detection module is relatively low.

Table 2
Average precision of detection computed on the IPHD-test dataset for Faster-RCNN after applying a priori (early fusion) and a posteriori (late fusion) merging techniques. Two best results – one for early fusion and one for late fusion are shown in bold print

Fusion strategy		Detection precision (std. dev.)		
		Thermal + depth images		
		AP_{50}	AP_{75}	MAP
Late fusion with Dual-NMS	Leave all unpaired	77.28% (±0.72)	55.14% (±0.48)	49.91% (±0.57)
	Leave unpaired thermal	**83.31% (±0.39)**	**57.84% (±0.14)**	**53.18% (±0.32)**
	Leave unpaired depth	74.90% (±0.56)	50.43% (±0.45)	47.07% (±0.54)
	Remove all unpaired	69.63% (±0.44)	52.98% (±0.14)	46.76% (±0.33)
Late fusion with Simple-NMS		73.26% (±0.38)	51.39% (±0.68)	46.62% (±0.47)
Early fusion		**88.86%** (±0.19)	**63.82%** (±0.21)	**57.42%** (±0.38)
Single model	Thermal	83.23% (±0.28)	56.38% (±0.32)	51.77% (±0.29)
	Depth	71.26% (±0.59)	44.01% (±0.45)	42.46% (±0.49)

4.3. Early Fusion

The early fusion approach requires modification of the model structure in order to accept concatenated data sources serving as input for the network. The proposed method encodes a pair of images consisting of one thermal image and one depth image in a single pass of the network. The images are rescaled to 1280×720 pixels to ensure their consistent size, and are then stacked to create a 2-channel multimodal image. As the model structure is less typical than the structures commonly used in Faster R-CNN networks, it limits the options of transfer learning and requires full retraining. The training dataset needs to be adequately prepared. In the experiments, both channels were standardized in the same way as in single-image detection. For a fair comparison, the fusion model was also realized using Faster R-CNN architecture with ResNet-50 backbone, as it was the case in the previous tests.

The early fusion approach surpassed all other methods tested in this paper. The precision rates obtained equaled $AP_{50} = 88.86\%$ and $AP_{75} = 63.82\%$, respectively. The result was better than for the single thermal model by 6.7% (AP_{50}) and 13.2% (AP_{75}), and surpassed the dual-NMS method by 6.6% (AP_{50}) and 10.3% (AP_{75}). Qualitative results for the early fusion model are shown in the third row of Fig. 5. Some of the missing and false detections were eliminated as a result of applying fusion to the detector.

5. Conclusions

The experiments described in this paper prove that it is possible to detect humans, both in thermal and depth images, while achieving acceptable precision rates. The rates are acceptable but still far from the precision level that may be obtained for RGB images. When comparing the two types of streams, it may be noticed that the images with information about the temperature are a better source of visual information for identifying people. In depth images, people are harder to distinguish from other objects, both for our algorithms and for humans. On the other hand, the thermal network has more false positives indicating other warm objects, such as animals or electric equipment. This is not the case in the depth model. We show that merging two types of imaging sources is capable of improving the outcomes generated by the detection network. We tested two important fusion strategies: early fusion (combining images at the network's input) and late fusion with a modified non-maximum suppression algorithm, namely dual-NMS. Both variants showed improvements in comparison to single-model detection. The best solution was obtained when the model was retrained on compound images (early fusion), and the result was better than the one achieved with the dual-NMS approach. However, we must stress that preparing early fusion models requires more effort and computational resources. We used an aligned dataset to make the training procedure possible. This type of data is not always available. Late fusion approaches allow merging the detection results from two independent single-source models trained on unaligned data. Further work on the methods under consideration requires the dataset to be extended. Access to an RGB source, for reference purposes, could be valuable for the development of the algorithm. The order of frames is another piece of information that is missing but may be obtained. After proper labeling, the detection method can be developed further and converted into a human tracking solution.

References

[1] A. Clapés, J. S. Jacques Junior, C. Morral, and S. Escalera, "Chalearn lap 2020 challenge on identity-preserved human detection: Dataset and results", in *15th IEEE Int. Conf. on Automatic Face and Gesture Recogn. (FG 2020)*, Buenos Aires, Argentina, pp. 859–866, 2020 (DOI: 10.1109/FG47880.2020.00135).

[2] T.-Y. Lin *et al.*, "Microsoft COCO: common objects in context", CoRR, abs/1405.0312, 2014 [Online]. Available: https://arxiv.org/pdf/1405.0312

[3] Teledyne Flir LLC, "LWIR micro thermal camera module Lepton 3", 2018 [Online]. Available: https://www.flir.com/products/lepton/?model=500-0276-01 (accessed on: 01.01.2021).

[4] J. Smisek, M. Jancosek, and T. Pajdla, *3D with Kinect*, *Consumer Depth Cameras for Computer Vision. Adv. in Computer Vision and Pattern Recogn.*, A. Fossati, J. Gall, H. Grabner, X. Ren, K. Konolige, Eds., pp. 3–25. Springer London, London: 2013 (DOI: 10.1007/978-1-4471-4640-7_1).

[5] Intel Corporation, Intel RealSense Depth Module D400 Series CustomCalibration, 2019 [Online]. Available: https://www.intel.com/content/dam/support/ us/en/documents/emerging-technologies/ intel-realsense-technology/RealSense_D400% 20_Custom_Calib_Paper.pdf (accessed on: 01.01.2021).

[6] S. Kumar, T. K. Marks, and M. Jones, "Improving person tracking using an inexpensive thermal infrared sensor", in *IEEE Conf. on Computer Vision and Pattern Recogn. Workshops*, Columbus, OH, USA, pp. 217–224, 2014 (DOI: 10.1109/CVPRW.2014.41).

[7] A. S. Charan, M. Jitesh, M. Chowdhury, and H. Venkataraman, "Abifn: Attention-based bi-modal fusion network for object detection at night time", *Electronics Letters*, vol. 56, no. 24, pp. 1309–1311, 2020 (DOI: 10.1049/el.2020.1952).

[8] H. Haggag, M. Hossny, S. Nahavandi, and O. Haggag, "An adaptable system for RGB-D based human body detection and pose estimation: Incorporating attached props", in *IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC)*, Budapest, Hungary, pp. 1544–1549, 2016 (DOI: 10.1109/SMC.2016.7844458).

[9] O. H. Jafari, D. Mitzel, and B. Leibe, "Real-time RGB-D based people detection and tracking for mobile robots and headworn cameras", in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, Hong Kong, China, 2014, pp. 5636–5643 (DOI: 10.1109/ICRA.2014.6907688).

[10] M. Rasoulidanesh, S. Yadav, S. Herath, Y. Vaghei, and S. Payandeh, "Deep attention models for human tracking using RGBD", *Sensors*, vol. 19, no. 4, 2019 (DOI: 10.3390/s19040750).

[11] H. S. Hadi, M. Rosbi, U. U. Sheikh, and S. H. M. Amin, "Fusion of thermal and depth images for occlusion handling for human detection from mobile robot", in *10th Asian Control Conf. (ASCC)*, Kota Kinabalu, Malaysia, pp. 1–5, 2015 (DOI: 10.1109/ASCC.2015.7244722).

[12] D. J. Yeong, G. Velasco-Hernandez, J. Barry, and J. Walsh, "Sensor and sensor fusion technology in autonomous vehicles: A review", *Sensors*, vol. 21, no. 6, 2021 (DOI: 10.3390/s21062140).

[13] V. F. Vidal *et al.*, "Sensors fusion and multidimensional point cloud analysis for electrical power system inspection", *Sensors*, vol. 20, no. 14, pp. 40–42, 2020 (DOI: 10.3390/s20144042).

[14] T. Alldieck, C. H. Bahnsen, and T. B. Moeslund, "Context-aware fusion of RGB and thermal imagery for traffic monitoring", *Sensors*, vol. 16, no. 11, 2016 (DOI: 10.3390/s16111947).

[15] F. Farahnakian and J. Heikkonen, "Deep learning based multi-modal fusion architectures for maritime vessel detection", *Remote Sensing*, vol. 12, no. 16, 2020 (DOI: 10.3390/rs12162509).

[16] A. Morfin-Santana *et al.*, "Real-time people detection from thermal images by using an unmanned aerial system", in *16th Int. Conf. on Electric. Engineer., Comput. Sci. and Automatic Control (CCE)*, Mexico City, Mexico, pp. 1–6, 2019 (DOI: 10.1109/ICEEE.2019.8884561).

[17] S. Chang, F. Yang, W. Wu, Y. Cho, and S. Chen, "Nighttime pedestrian detection using thermal imaging based on hog feature", in *Proc. 2011 Int. Conf. on System Sci. and Engineer.*, Macau, China, pp. 694–698, 2011 (DOI: 10.1109/ICSSE.2011.5961992).

[18] L. Spinello and K. O. Arras, "People detection in RGB-D data", in *IEEE/RSJ Int. Conf. on Intell. Robots and Systems*, San Francisco, CA, USA, pp. 3838–3843, 2011 (DOI: 10.1109/IROS.2011.6095074).

[19] Ch. Herrmann, M. Ruf, and J. Beyerer, "CNN-based thermal infrared person detection by domain adaptation", in *Proc. Autonomous Systems: Sensors, Vehicles, Security, and the Internet of Everything*, M. C. Dudzik and J. C. Ricklin, Eds., Orlando, FL, USA, vol. 10643, 2018, pp. 38–43 (DOI: 10.1117/12.2304400).

[20] Sh. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks", in *Proc. of the 28th Int. Conf. on Neural Information Process. Systems*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, Eds., vol. 1, pp. 91–99, 2015 [Online]. Available: http://papers.nips.cc/paper/5638-faster-r-cnn-towards-real-time-object-detection-with-region-proposal-networks.pdf

[21] K. He, G. Gkioxari, P. Dollár, and R. B. Girshick, "Mask R-CNN", CoRR, abs/1703.06870, 2017 [Online]. Available: https://arxiv.org/pdf/1703.06870

[22] W. Liu *et al.*, "SSD: single shot multibox detector", CoRR, abs/1512.02325, 2015 [Online]. Available: https://arxiv.org/pdf/1512.02325

[23] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement", CoRR, abs/1804.02767, 2018 [Online]. Available: https://arxiv.org/pdf/1804.02767

[24] K. He, X. Zhang, Sh. Ren, and J. Sun, "Deep residual learning for image recognition", CoRR, abs/1512.03385, 2015 [Online]. Available: https://arxiv.org/pdf/1512.03385

[25] A. Karpathy *et al.*, "Large-scale video classification with convolutional neural networks", *27th IEEE Conf. on Computer Vision and Pattern Recogn. (CVPR)*, Columbus, OH, USA, 2014 (DOI: 10.1109/CVPR.2014.223).

[26] A. Neubeck and L. Van Gool, "Efficient non-maximum suppression", in *18th Int. Conf. on Pattern Recogn. (ICPR'06)*, Hong Kong, China, vol. 3, pp. 850–855, 2006 (DOI: 10.1109/ICPR.2006.479).



Weronika Gutfeter is a research assistant at the Biometric and Machine Intelligence Laboratory of the Research and Academic Computer Network (NASK) in Warsaw, Poland. She obtained her Master's degree in Computer Engineering from Warsaw University of Technology and is currently working on her Ph.D. thesis that covers the topics of 3D, semi-3D and multi-view face identification methods. Her research interests include also other computer vision and biometric identification methods, such as off-line signature verification and iris detection. At NASK, she works on developing machine learning systems for the public sector.

https://orcid.org/0000-0001-6359-8220
E-mail: weronika.gutfeter@nask.pl
Biometric and Machine Intelligence Laboratory
Research and Academic Computer Network (NASK)
ul. Kolska 12
Warsaw, Poland



Andrzej Pacut is a full Professor, chair of the Biometrics and Machine Learning Group at the Faculty of Electronics and Information Technology in Warsaw University of Technology. He is the founder of the Technical Committee on Biometrics (mirror of ISO/IEC JTC1 SC37) within the Polish Standardization Committee (PKN), chair of the Biometrics and Machine Intelligence Laboratory at the Research and Academic Computer Network (NASK), and a member of the NASK Research Council.

https://orcid.org/0000-0003-3489-8990
E-mail: andrzej.pacut@nask.pl
Biometric and Machine Intelligence Laboratory
Research and Academic Computer Network (NASK)
ul. Kolska 12
Warsaw, Poland

Paper

Has the Internet Saved the Economy? Modeling Impact of ICT Sector and COVID-19 on GDP

Magdalena Olender-Skorek, Michał Szałański, and Marek Sylwestrzak

National Institute of Telecommunications, Warsaw, Poland

https://doi.org/10.26636/jtit.2021.157821

Abstract—This paper presents the influence of the COVID-19 pandemic on gross domestic product (GDP) per capita for the 27 countries of the European Union. A panel model with fixed effects was applied to a dataset from 2010 to 2020. The analysis covered 13 independent variables, including nine related to the telecommunications market, and assessed their impact on GDP per capita. A variable related to the number of COVID-19 deaths per one thousand inhabitants was then added to the model. The results showed that COVID-19 is a significant factor and is negatively correlated with GDP per capita. The analysis described in the article has also shown that the importance of the ICT sector increased during the pandemic, i.e. the household broadband Internet variable became statistically significant.

Keywords—COVID-19, European Union, GDP, ICT sector.

1. Introduction

The impact that technology exerts on the economy is indisputable, as evidenced by theoretical and empirical research conducted in this area. Nowadays, innovative technologies play a very significant role: they increase productivity, redefine the manufacturing paradigms, remodel supply chain relations and influence consumption [1]–[3].

In this paper, we examine how the relationship between the market of information and communication technologies (ICT) and the economy was affected by the COVID-19 pandemic. An analysis of the increasing number of publications indicates that modern technologies help mitigate the outcomes of the pandemic. In fact, studies conducted by international organizations suggest that countries with high technological potential cope with the pandemic crisis better than others [4]–[5].

Singh and Garg have also indicated that the telecommunications industry, in response to the changes in customer demand for telecommunication services resulting from COVID-19, must offer new digital products and tools and has to upgrade its network infrastructure due to increasing network traffic [6].

In contrast, Sale, Wood and Rebbeck show that revenues of telcos in developed countries declined by 3.4% in 2020 compared to 2019. This is due to reduced activity of telcos' customers – a phenomenon triggered by increased unemployment and economic slowdown [7].

Do the conclusions drawn from global studies apply to the more homogeneous structure of the 27 European Union (EU) countries? Has the pandemic in the EU affected the ICT market and economy? Is the ICT market affecting the economy with the same strength as it did before the pandemic? In this paper, authors will attempt to provide answers to those questions.

In this article, an econometric model will be constructed to verify whether there is a statistically significant relationship between the ICT market and gross domestic product per capita in the EU. The model will also examine the potential impact of COVID-19, measured as a number of deaths per 1,000 inhabitants.

So far, the impact of COVID-19 on the global economy and ICT market has not been the subject of extensive research. Only the International Telecommunication Union (ITU) has dealt with this issue extensively [4]–[5].

There have also been a few unrelated publications on the changes that the pandemic caused in the telecommunications market, but they focused on specific issues, e.g. on the increase in network traffic [8].

Several other studies devoted to regulations and the need for post-pandemic changes may also be identified. Research conducted in relation to this paper fills a certain gap in the existing work on EU countries and the impact of COVID-19 on their economies.

2. Theoretical Framework

The role of technological progress in the economy, presented from the point of view of different economic theories, serves as the point of departure for the considerations presented in this article. Solow proposed his model of economic growth by introducing a technological factor [9], [10]. The production function has the following form:

$$Y\ (t)\ =\ A(t)F[K(t),\ L(t)]\ , \tag{1}$$

where t is the year, $Y(t)$ is total production, $A(t)$ is technological progress, $K(t)$ is capital accumulation, $L(t)$ is labor force, and $F[K(t),\ L(t)]$ is the production function.
The Cobb-Douglas function is most commonly taken here as the production function, hence the production equation takes the following form:

$$Y(t) = AK(t)^{\alpha}\ L(t)^{\beta},\ \alpha,\beta > 0,\ \alpha+\beta = 1\ , \qquad (2)$$

where α is elasticity of capital and β is elasticity of production.
Variable A is sometimes introduced endogenously into the production function as technical progress embedded in labor (also the so-called labor-augmenting technical progress) or in capital (the so-called capital-augmenting technical progress). Then, the formula of such a production function is:

$$Y(t) = K(t)^{\alpha}AL(t)^{\beta}\ , \qquad (3)$$

or, respectively:

$$Y(t) = aK(t)^{\alpha}L(t)^{\beta}\ . \qquad (4)$$

However, due to the key importance of ICT, growth models that explicitly include variables related to this sector are becoming increasingly popular in the literature. Researchers test the importance of the ICT market for the economy by measuring its influence on GDP and productivity, but also test the impact of regulations on investment in new technologies. The following function is an example of such an approach [10]:

$$Y = AL^{\alpha}(N-ICT)^{\beta}ICT^{\gamma}\ , \qquad (5)$$

where: Y is the denotation of a country's GDP, ICT is the capital spent on ICT, N-ICT is the remaining capital (non-ICT), α – is elasticity of production, and β – is elasticity of non-ITC capital.
Modeling for Poland was carried out by Kaczmarczyk [11], who used 16 variables describing the ICT sector in detail and several control macroeconomic variables, in his study of the interaction between ICT and GDP. The model originally included, inter alia, the number of people employed in the ICT sector, the number of ICT entrepreneurs (production and services), the value of net sales (production and services), the number of employees in ICT (production and services), and the research and development (R&D) expenditure.
Unlike the majority of other approaches, the author used net sales of the ICT sector as the explanatory variable. He then attempted to explain it mainly by R&D expenditure, because this variable was the only one that was statistically significant for ICT.
Models created in DELab UW are closer to their theoretical counterparts [12]. They rely on both general macroeconomic variables and factors that are directly related to the ICT sector to estimate the relationship between output (GDP per capita) and the ICT market. In their analyses of the latter variables, the authors selected 53 indicators, some of which proved to be statistically insignificant. Starting directly from the Solow model, the authors derived the following function:

$$\ln(GDP)_{it} = \ \beta_0 + (\beta_1+1)\ \ln GDP_{i,t-1} + x'_{it}\beta + \alpha_i + \varepsilon_{it}\,, \qquad (6)$$

where: GDP_{it} is gross domestic product in year t for country i, $\mathrm{GDP}_{i,t-1}$ is gross domestic product in year $t-1$ for country i, x'_{it} is vector of explanatory variables, α_i is the country-specific effect, ε_{it} is the random component.
The inclusion of GDP in period $t-1$ is explained by income level convergence (i.e. countries with a higher baseline GDP per capita experiencing lower GDP growth). The individual effect for a given country results, in turn, from the high diversity of the countries analyzed (the model considers data from Europe, Asia and Africa).
Recent econometric models, including those attemting to account for effects of the pandemic, are based on a similar convention. ITU attempts to estimate the relationship between ICT and the economy by designing a structural model in which the first equation takes the form of [4]:

$$\log(GDP)_{pc\,i} = \mu_i + \theta\log(GFKF)_{it} + \sigma\log(HK)_{it}$$
$$+\beta\log(BB\ PEN)_{it} + \delta COVID_{it} + \gamma(BB\ PEN\cdot COVID^2)_{it}$$
$$+\rho_{i,2020} + \tau_t + \varepsilon_{it}\ , \qquad (7)$$

where: GDP_{pc} is Gross Domestic Product per capita, GFKF stands for gross fixed capital formation, HK is human capital, BB PEN is broadband penetration, $\rho_{i,2020}$ is the individual country effect, τ_t are the control variables, COVID is the number of deaths from SARS-CoV-2 per 100 people. In addition, ITU measures the impact of digitization on the economy based on a variable called *digitization index*, as well as other factors: labor force, capital, and previous year's GDP. The results of the study showed that in countries with a more developed ICT market, the impact of the pandemic on GDP was lower. This means that the ICT market remains an important factor in the development of the economy, even during a pandemic.

3. Data – Empirical Specification

Development of the following models was based on the Solow model, with later modifications, while the selection of variables was determined by econometric studies performed around the world and by data availability. The data used for this analysis are sourced from Eurostat, the World Bank, and ITU databases. The modeling considered annual data for the 27 EU countries, from the period between 2010 and 2020 decade. The variables that were used in the models are presented in Table 1.
The modeling started with the analysis of raw data and of the GDP per capita variable mentioned above, and relied on the histograms of its individual determinants. In the

Table 1
Variables used for modeling; source: authors' analysis

Symbol	Variable	Source
gdp	GDP per capita in current prices	Eurostat
dgdp	Lagged GDP per capita in current prices	Eurostat
employment	Total employment from 20 to 64 years [in thousands]	Eurostat
employment_ICT	Total employment in ICT sector [in thousand]	Eurostat
r_d	Total government budget allocations for R&D [in millions of euros]	Eurostat
dgov_expenditure	Total general government expenditure [in millions of euros]	Eurostat
dimport_ICT	ICT goods imports as percentage of total goods imports	World Bank
dexport_ICT	ICT goods exports as percentage of total goods exports	World Bank
household_broadband	Household broadband Internet connection [percentage]	Eurostat
household_access	Households – level of Internet access [percentage]	Eurostat
household_fixed	Fixed broadband subscriptions per 100 inhabitants	ITU
household_mobile	Active mobile-broadband subscriptions per 100 inhabitants	ITU
price_fixed	Fixed broadband basket	ITU
price_mobile	Mobile broadband basket	ITU
covid	COVID death per 1000 inhabitants	John Hopkins University

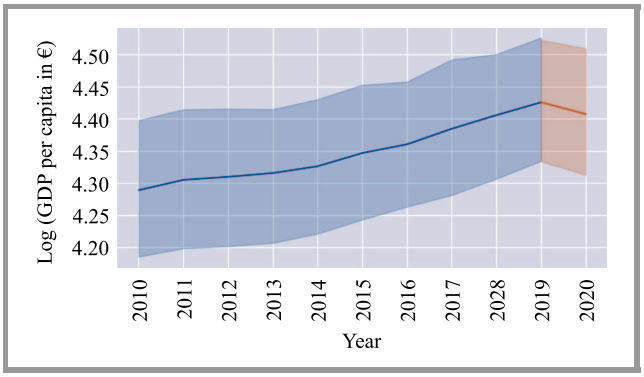


Fig. 1. GDP per capita in EU countries in 2010–2020. Source: authors' analysis.

case of GDP per capita, even a simple comparison with individual countries, through the years, shows that a clear decrease was observed in 2020 (Fig. 1).

Apart from the pandemic, no other exogenous factors affecting GDP were observed at that time. This suggests that it was COVID-19 that has led to the decrease in GDP per capita [13]. Lockdowns, health care problems, production shutdowns, electronic chip shortages or restrictions affecting the transportation of goods and services are just some of the factors that directly affect the economy.

The Hadri test for unit roots in panel data returned a p-value of 0.000. This means that we can reject the null hypothesis that no series has a unit root. By further testing each country separately, using the augmented Dickey-Fuller test (p-value threshold of 0.05), we conclude that all series are non-stationary, with the exception of Greece, Luxembourg, Slovenia, Finland, and Sweden.

To account for the 2020 pandemic, we included a control variable for the number of COVID-19 deaths per 1,000 people. The hypothesis is that with a greater number of infected people, the isolation measures enacted would be broader and harsher. This, in turn, could harm the economy. The number of deaths was chosen because the number of infections does not account for different testing strategies employed by various countries in response to the pandemic. The distributions of the individual variables were close to normal distribution, with some of them seeming to be left skewed. On the other hand, correlations between variables turned out to be strong for ICT imports and exports, as well as for GDP per capita and lagged GDP per capita. Detailed results of the correlation analysis performed are presented in Fig. 2.

4. Econometric Model Results

The models that were tested for the purpose of this analysis were based on PooledOLS, random effects models, and fixed effects models [14], [15]. The F-tests for poolability had a p-value of 0.000 for the pre-Covid era and of 0.000 with the year 2020 included. This means that we can reject the null hypothesis that the countries are homogeneous. Hence, the PooledOLS model is not a good fit for this analysis, as it does not account for the individual effects [16]–[19].

Residuals of the PooledOLS model were also tested using the Ljung-Box and Box-Pierce tests for autocorrelation of the residuals. The resulting p-values were 0.000 and 0.000, respectively. Therefore, we can reject the null hypothesis that the results are not autocorrelated. This indicates that

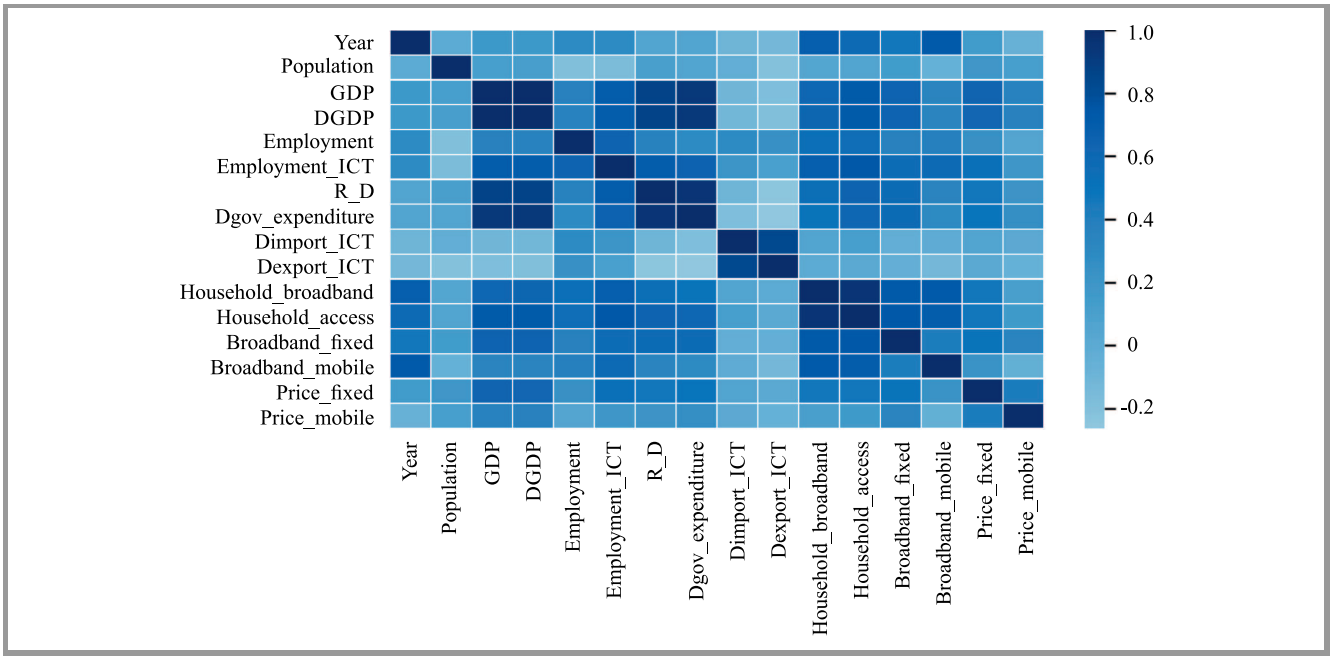


Fig. 2. Correlation between independent variables. Source: authors' analysis.

the residuals are autocorrelated, which is a second reason that the PooledOLS model is not suitable for this analysis.

The decision was made to use the Durbin-Wu-Hausman tests. The p-values were: 0.000 for the pre-Covid era and 0.000 with 2020 included. This means that we reject the null hypothesis that the random effects model is consistent, and thus the fixed effects model is preferred, because it is still consistent under the alternative hypothesis.

Estimation of the fixed effects model for 2010-2019 indicated 5 statistically significant variables, of which only the amount of government expenditure negatively affected the dependent variable. The others, i.e. previous year's GDP, total and ICT employment, and households' access to the Internet affected the GDP positively (Table 2, FE pre-Covid column).

Including year 2020 in the analysis and adding the Covid variable reduced the fit of the model but made the variable representing the price of landline calls statistically significant (Table 2, FE Covid model column).

After estimating the preliminary results, only those variables that were found to be statistically significant were left for the final analysis, with p-values equal to or less than 0.05 (Table 2, FE Covid optimized model column).

The residuals of the *FE Covid optimized* model were tested using the Breusch Pagan test for heteroscedasticity. The p-value was 0.2398, therefore there is no need to reject the null hypothesis that the residuals of the model are homoscedastic [20], [21].

Since the Durbin Watson test for autocorrelation cannot be used, as the lagged dependent variable is used as an independent one, autocorrelation was tested using the Ljung-Box test. The resulting p-value for lag = 1 is 0.2081. Therefore, there is no point to reject the null hypothesis that the residuals are not autocorrelated. The similar Box-Pierce test resulted in the p-value of 0.2104, with the same outcome.

The residuals follow a normal distribution, with 2 exceptions (Fig. 3). The residual on the right is Ireland in 2015, where the GDP rose by 26%. The residuals on the extreme left are a few countries in 2020. Each country handled the COVID-19 outbreak differently, and this variance was not fully explained by the number of deaths per capita.

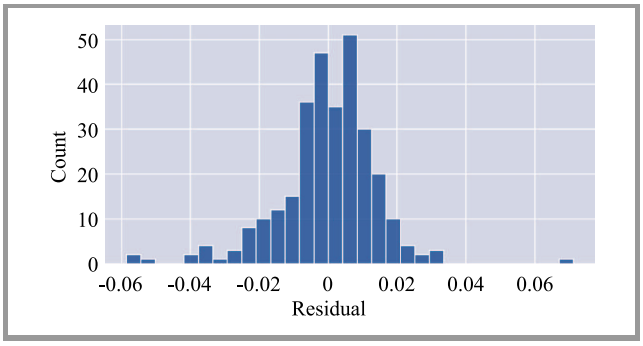


Fig. 3. Distributions of residuals in the FE Covid optimized model.

The assumptions of the fixed effects model residuals are as follows: they have a mean of 0, they are normally distributed and their variance is constant. They are also not autocorrelated. The residuals of the *FE Covid optimized* model meet these assumptions.

Finally, it was decided to use the model in its general form:

$$y_{it} = \beta_0 + \beta X_{it} + \alpha_i + u_{it} \tag{8}$$

for $t = 1, \ldots, T$ and $i = 1, \ldots, N$, where: β_0 is the constant, β is the matrix of estimated parameters, X_{it} is the

Table 2
Estimation results

Symbol	FE pre-Covid	FE Covid	FE Covid optimized
const	0.7634 (0.0000)[1)]	0.7496 (0.0000)[1)]	0.6769 (0.0000)[1)]
dgdp	0.7870 (0.0000)[1)]	0.7981 (0.0000)[1)]	0.8182 (0.0000)[1)]
employment	0.3341 (0.0001)[2)]	0.2886 (0.0006)[2)]	0.2896 (0.0001)
employment_ICT	1.3542 (0.0131)	1.1061 (0.0460)	
r_d	-0.0000 (0.7802)	-0.0000 (0.3429)	
dgov_expenditure	-0.0000 (0.0000)[1)]	-0.0000 (0.0001)[1)]	-0.0000 (0.0000)[1)]
dimport_ICT	0.0010 (0.3836)	0.0008 (0.4808)	
dexport_ICT	-0.0005 (0.4215)	-0.0002 (0.7656)	
household_broadband	0.0003 (0.2461)	0.0006 (0.0298)[2)]	0.0023 (0.0023)[1)]
household_access	0.0004 (0.2504)	-0.0000 (0.8122)	
household_fixed	-0.0000 (0.8828)	-0.0001 (0.8332)	
household_mobile	0.0001 (0.0480)[2)]	0.0001 (0.0392)[2)]	0.0002 (0.0025)[1)]
price_fixed	0.0002 (0.1739)	0.0003 (0.0242)[2)]	0.0003 (0.0088)[1)]
price_mobile	-0.0000 (0.9062)	0.0000 (0.9156)	
covid		-0.0330 (0.0000)[1)]	-0.0325 (0.0000)[1)]
observations	270	297	297
number of countries	27	27	27
R-squared	0.9576	0.9515	0.9276
[1)], [2)], [3)] significant at 1 per cent, 5 per cent and 10 per cent critical value, respectively. Note: dgdp is expressed as a decimal logarithm.			

time-variant vector containing independent variables, α_i is the time-invariant, unobserved effect of each individual entity, and u_{it} is the error term.

After filling it in with data, Eq. (8) took the form:

$$\begin{aligned}\log(gdp)_{it} &= 0.6769 + 0.8182\log(gdp)_{i,t-1} \\ &+0.2896\ employment_{it} - 0.000004303\ gov.expenditure_{it} \\ &+0.0005\ household.broadband_{it} \\ &+0.0002\ broadband.mobile_{it} + 0.0003\ price.fixed_{it} \\ &-0.0325 covid_{it} + \alpha_i + u_{it}\end{aligned} \tag{9}$$

for $t = 1, \ldots, 11$ and $i = 1, \ldots, 27$.

4.1. Interpretation of Model Results

We find R-squared value of 0.9276 to be a good result indicating the models may be used to examine the relationship between the economy, the ICT market and COVID-19. All variables in the final *FE Covid optimized* model are significant with p-values of less than 0.01.

From the resulting model it was obtained that, ceteris paribus, if the previous year's GDP increased by 1%, then the current year's GDP would increase by 0.7981%. Ceteris paribus, an increase in employment in the economy by 1% would result in an increase in GDP by 0.336%. Smaller differences can be observed for other variables related to the ICT market, the change of which would result in an increase in GDP ranging from 0.02% to 0.23% (ceteris paribus). On the other hand, an increase in COVID-related deaths by 1 per 1000 inhabitants would decrease GDP by 0.032% (ceteris paribus).

The *FE Covid optimized* model seems to confirm the hypothesis that there could be a causal link between the number of deaths due to COVID-19 and GDP dynamics. The correlation is significant, the time sequence is correct, i.e. the causes precede the consequences, and there exists a plausible explanation of the relationship between the variables.

However, the model exhibits some outliers in 2020. This means that the COVID-19 death variable alone is not enough to account for the unobserved, time-variant variable that represents the isolation measures introduced.

5. Further Research Areas

Despite the good fit of the model, it is not free of flaws and limitations. First, the studied group of countries is not homogeneous – both geopolitically and in terms of economic development. The formal regulations existing in the ICT sector and the level of awareness of modern technologies differ as well. These factors were not considered directly in this analysis, but a model was used that takes such diversity into account using the time-invariant unobserved effects of the individual entities. An additional rule used to offset this heterogeneity is the introduction, to the model of the lagged GDP per capita variable. Admittedly, it explains a large portion of the explanatory variable variance, but this makes the impact of other variables more comparable and closer to reality.

Some unusual observations were made concerning the raw data, showing that four countries deviate significantly from the remaining group. The degree of these deviations is large enough to consider removing them from further analysis, which could further improve the model's fit. The observed anomalies are:

- Ireland, which experienced record high GDP growth of over 26% in 2015 (explained in the literature as an accounting effect),
- Cyprus and Malta – which are small island countries, hardly comparable in terms of ICT develop-

ment, clearly different from the other countries under analysis,

- Sweden – where the residuals of the model strongly deviate from the residuals of the model for other countries.

A certain weakness of the model and the analysis performed may also be that a small number of independent variables was examined. However, this is a limitation that is difficult to overcome, because it is external and results from the changing data collection methodology and the lack of available comparable statistics for different countries. Due to the panel nature of the data, any single deficiency in the statistics methodology results in the exclusion of individual variables from the model.

Another limitation of the model is the relatively small number of observations that include the COVID-19 pandemic. This variable (although statistically significant) appears only in 2020. Inclusion of the pandemic in the following year may affect the results of the model. It is expected to strengthen the effect of the pandemic and the ICT market itself. The estimation is undoubtedly worth repeating, but also (with enough data) one might be tempted to create a separate model exclusively for data from the pandemic period, and to compare its results with the model from the period before the pandemic (up to 2019).

It is worth paying attention to the COVID-19 variable itself and introducing it into the model in a modified form, not only as the number of deaths per 1,000 inhabitants. It turns out that in this form it does not explain all the variability. Other variables representing the economic impact (number and severity of lockdowns, travel bans, etc.), would certainly help explain more of the variances.

It is expected that the results of further, in-depth research will not substantially change the basic conclusions of the presented analysis. However, few details of the model may be modified, and some variables that were initially insignificant may finally become statistically significant.

6. Conclusions

Based on the study which covered 27 EU countries, the statistical impact of COVID-19 on the economy may be confirmed. Firstly, it was shown that the COVID-19 variable is statistically significant. Secondly, and more importantly, the results obtained show that with the outbreak of the pandemic, the ICT market has gained in importance. The previously statistically insignificant variable of household broadband Internet penetration has become statistically significant.

Real world observations lead us to conclusions that are similar to the model results - the introduction of lockdowns forced households to work and study remotely. Network traffic was not the only metric that increased. So did the quality of service requirements. The need to handle larger data volumes with minimal latency has emerged, and broadband connectivity is the answer here.

This study has shown that differences between EU countries do matter and cannot be ignored in the modeling. A comparison of the model that treats the EU countries as a relatively homogeneous structure with models that take into account variations existing between countries (by introducing additional variables for each country) clearly favors the latter. This observation is consistent with the results of previous market research [12].

References

[1] M. Olender-Skorek, "Czwarta rewolucja przemysłowa a wybrane aspekty teorii ekonomii", *Nierówności Społeczne a Wzrost Gospodarczy*, vol. 51, pp. 38–49, 2017 (DOI: 10.15584/nsawg.2017.3.3) [in Polish].

[2] E. Mączyńska and E. Okoń-Horodyńska, "Przedsiębiorstwo i jego otoczenie w obliczu czwartej rewolucji przemysłowej – Wyzwania, szanse i zagrożenia", *Przegląd organizacji*, vol. 1, pp. 9–21, 2020 (DOI: 10.33141/po.2020.01.01) [in Polish].

[3] T. Zalega, "Nowe trendy konsumenckie jako przejaw innowacyjnych zachowań współczesnych konsumentów", *Nierówności Społeczne a Wzrost Gospodarczy*, vol. 46, pp. 202–225, 2016 (DOI: 10.15584/nsawg.2016.2.11) [in Polish].

[4] ITU, "The economic impact of broadband and digitization through the COVID-19 pandemic. Econometric modelling", 2021 [Online]. Available: https://www.itu.int/dms_pub/itu-d/opb/pref/D-PREF-EF.COV_ECO_IMPACT_B-2021-PDF-E.pdf

[5] ITU, "How broadband, digitization and ICT regulation impact the global economy. Global econometric modeling", 2020 [Online]. Available: https://www.itu.int/dms_pub/itu-d/opb/pref/D-PREF-EF.BDR-2020-PDF-E.pdf

[6] G. Singh and S. Garg, "Affect of Covid-19 on telecom sector in India – A review", *PalArch's J. of Archaeology of Egypt/Egyptology*, vol. 17, no. 6, pp. 13206–13213, 2020 [Online]. Available: https://archives.palarch.nl/index.php/jae/article/download/3399/3390/6555

[7] S. Sale, R. Wood, and T. Rebbeck, "COVID-19 will lead telecoms revenue to decline by 3.4% in developed markets in 2020", *Analysys Mason Report*, 2020 [Online]. Available: https://www.analysysmason.com/research/content/short-reports/covid-19-operator-revenue-impact/

[8] M. Karpowicz, "COVID-19 pandemic and Internet traffic in Poland: evidence from selected regional networks", *J. of Telecommun. and Informat. Technol.*, vol. 3, pp. 86–91, 2021 (DOI: 10.26636/jtit.2021.15472).

[9] R. Solow, "A contribution to the theory of economic growth", *The Quarterly J. of Economics*, vol. 70, no. 1, pp. 65–94, 1956 (DOI: 10.2307/1884513).

[10] I. Telega, "Trwałość w modelu wzrostu Solowa. Analiza krytyczna", *Zeszyty Naukowe/Polskie Towarzystwo Ekonomiczne*, vol. 12, pp. 21–33, 2012 [Online]. Available: http://www.krakow.pte.pl/pliki/zn-pte-nr-12/zn-12telega.pdf [in Polish].

[11] P. Kaczmarczyk, "Econometric modelling and forecasting of the Polish ICT sector development with regard to macroeconomic indicators", *Roczniki Kolegium Analiz Ekonomicznych*, vol. 45, pp. 259–272, 2017 [Online]. Available: http://rocznikikae.sgh.waw.pl/p/roczniki_kae_z45_19.pdf [in Polish].

[12] M. Próchniak and B. Witkowski, "Digitalizacja i internetyzacja a wzrost gospodarczy", *Working Paper DELab UW*, vol. 4, 2016 [Online]. Available: https://www.delab.uw.edu.pl/wp-content/uploads/2020/07/digitalizacja_i_internetyzacja.pdf [in Polish].

[13] M. König and A. Winkler, "COVID-19: Lockdowns, Fatality Rates and GDP Growth Evidence for the First Three Quarters of 2020", *Intereconomics*, vol. 56, no. 1, pp. 32–39, 2021 (DOI: 10.1007/s10272-021-0948-y).

[14] B. Baltagi, *Econometric Analysis of Panel Data*, Hoboken: John Wiley & Sons, 2008 (ISBN: 9783030539535).

[15] W. Greene, *Econometric Analysis*, New Jersey: Pearson Education, 2003 (ISBN: 9780130661890).

[16] S. Seabold and J. Perktold, "Statsmodels: Econometric and statistical modeling with Python", *Proc. of the 9th Python in Sci. Conf. SciPy*, Austin, TX, USA, 2010, vol. 57, pp. 92–96 [Online]. Available: http://conference.scipy.org/proceedings/scipy2010/pdfs/seabold.pdf

[17] C. Harris *et al.*, "Array programming with NumPy", *Nature*, vol. 585, pp. 357–362, 2020 (DOI: 10.1038/s41586-020-2649-2).

[18] W. McKinney, "Data structures for statistical computing in Python", *Proc. of the 9th Python in Sci. Conf. SciPy*, 2010, vol. 445, pp. 56–61 [Online]. Available: http://conference.scipy.org/proceedings/scipy2010/pdfs/mckinney.pdf

[19] R. Zbyrowski, "Szacowanie wartości nieruchomości na podstawie modeli ekonometrycznych", *Equilibrium. Quarterly J. of Economics and Economic Policy*, vol. 4, no. 1, pp. 241–252, 2010 (DOI: 10.12775/EQUIL.2010.019).

[20] J. Hunter, "Matplotlib: A 2D graphics environment", *Comput. in Sci. & Engineer.*, vol. 9, no. 3, pp. 90–95, 2007 (DOI: 10.1109/MCSE.2007.55).

[21] M. Waskom, "Seaborn: statistical data visualization", *J. of Open Source Software*, vol. 6, no. 60, pp. 1–4, 2021 (DOI: 10.21105/joss.03021).



Magdalena Olender-Skorek holds a Ph.D. in Economics and is an Assistant Professor at the Market Economy Unit of the Department of National Economy, Faculty of Management, University of Warsaw. She is also an Assistant Professor at the National Institute of Telecommunications. She is a member of the Centre for Antitrust and Regulatory Studies (CARS), a member of INE PAN Team drawing up the "Polish Economy – Forecasts and Opinions Report" (2006–2013), as well as an academic lecturer and coach. Her interests focus on the telecom market, competition policy, market regulations, and infrastructure industries. She is an author and co-author of numerous expert opinions, reports, articles and chapters of joint publications.

https://orcid.org/0000-0002-4831-6122
E-mail: M.Olender@il-pib.pl
National Institute of Telecommunications
Warsaw, Poland



Michał Szałański holds a B.Sc. in Finance and is a Data Scientist at the National Institute of Telecommunications. He is currently completing his M.Sc. (Data Science) at the Faculty of Economic Sciences of the University of Warsaw. He is responsible for developing supervised and unsupervised machine learning models used to tackle data compatibility problems in telemetry, using state-of-the-art methods to merge hundreds of millions of datapoints. Results of his work were presented at the prestigious Economic Forum in Krynica, Poland.

https://orcid.org/0000-0002-3131-6193
E-mail: M.Szalanski@il-pib.pl
National Institute of Telecommunications
Warsaw, Poland



Marek Sylwestrzak received his M.Sc. and Ph.D. degrees from the Faculty of Economic Sciences, Warsaw University, in 2010 and 2018, respectively. Since 2010, he has been working at the ICT Market Analysis and Development Department of the National Institute of Telecommunications. He is an author and co-author of numerous publications. His research interests focus on the ICT and analysis of economic influence macroeconomic indicators.

https://orcid.org/0000-0001-8962-8168
E-mail: M.Sylwestrzak@il-pib.pl
National Institute of Telecommunications
Warsaw, Poland

Information for Authors

Journal of Telecommunications and Information Technology **(JTIT)** is published quarterly. It comprises original contributions, dealing with a wide range of topics related to telecommunications and information technology. **All papers are subject to peer review**. Topics presented in the JTIT report primary and/or experimental research results, which advance the base of scientific and technological knowledge about telecommunications and information technology.

JTIT is dedicated to publishing research results which advance the level of current research or add to the understanding of problems related to modulation and signal design, wireless communications, optical communications and photonic systems, voice communications devices, image and signal processing, transmission systems, network architecture, coding and communication theory, as well as information technology.

Suitable research-related papers should hold the potential to advance the technological base of telecommunications and information technology. Tutorial and review papers are published only by invitation.

Manuscript. TEX and LATEX are preferable, standard Microsoft Word format (.doc) is acceptable. The authors JTIT LATEX style file is available:

https://www.itl.waw.pl/en/jtit-for-authors

Papers published should contain up to 10 printed pages in LATEX authors style (Word processor one printed page corresponds approximately to 6000 characters).

The manuscript should include an abstract about 150–200 words long and the relevant keywords. The abstract should contain statement of the problem, assumptions and methodology, results and conclusion or discussion on the importance of the results. Abstracts must not include mathematical expressions or bibliographic references.

Keywords should not repeat the title of the manuscript. About four keywords or phrases in alphabetical order should be used, separated by commas.

The original files accompanied with pdf file should be submitted by e-mail: redakcja@itl.waw.pl

Figures, tables and photographs. Original figures should be submitted. Drawings in Corel Draw and PostScript formats are preferred. Figure captions should be placed below the figures and can not be included as a part of the figure. Each figure should be submitted as a separated graphic file, in .cdr, .eps, .ps, .png or .tif format. Tables and figures should be numbered consecutively with Arabic numerals.

Each photograph with minimum 300 dpi resolution should be delivered in electronic formats (TIFF, JPG or PNG) as a separated file.

References. All references should be marked in the text by Arabic numerals in square brackets and listed at the end of the paper in order of their appearance in the text, including exclusively publications cited inside. Samples of correct formats for various types of references are presented below:

[1] Y. Namihira, Relationship between nonlinear effective area and mode field diameter for dispersion shifted fibres, *Electron. Lett*., vol. 30, no. 3, pp. 262–264, 1994.

[2] C. Kittel, *Introduction to Solid State Physics*. New York: Wiley, 1986.

[3] S. Demri and E. Orłowska, Informational representability: Abstract models versus concrete models, in *Fuzzy Sets, Logics and Knowledge-Based Reasoning*, D. Dubois and H. Prade, Eds. Dordrecht: Kluwer, 1999, pp. 301–314

Biographies and photographs of authors. A brief professional authors biography of up to 200 words and a photo of each author should be included with the manuscript.

Galley proofs. Authors should return proofs as a list of corrections as soon as possible. In other cases, the article will be proof-read against manuscript by the editor and printed without the author's corrections. Remarks to the errata should be provided within one week after receiving the offprint.

Copyright. Manuscript submitted to JTIT should not be published or simultaneously submitted for publication elsewhere. By submitting a manuscript, the author(s) agree to automatically transfer the copyright for their article to the publisher, if and when the article is accepted for publication. The copyright comprises the exclusive rights to reproduce and distribute the article, including reprints and all translation rights. No part of the present JTIT should not be reproduced in any form nor transmitted or translated into a machine language without prior written consent of the publisher.

For copyright form see:
https://www.itl.waw.pl/en/jtit-for-authors

Journal of Telecommunications and Information Technology has entered into an electronic licencing relationship with EBSCO Publishing, the worlds most prolific aggregator of full text journals, magazines and other sources. The text of *Journal of Telecommunications and Information Technology* can be found on EBSCO Publishings databases. For more information on EBSCO Publishing, please visit www.epnet.com.

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY 4/2021

National Institute
of Telecommunications

Editorial Office

National Institute
of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

tel. +48 22 512 81 83
fax: +48 22 512 84 00
e-mail: redakcja@itl.waw.pl
http://www.nit.eu