

The Unwanted Emission Signals in the Context of the Reconstruct Possibility of Data Graphics

Ireneusz Kubiak

Military Communication Institute, Zegrze, 05-130, Poland
Email: i.kubiak@wil.waw.pl

Abstract—In its operation, every electrical device generates electromagnetic disturbance signals. They can be due to the operation of components of the device (step motors, heaters, control circuits, or electronic circuits). Quite often such signals have the characteristics of the data processed on such devices. They can have the form of a text. In each case, such signals are undesirable. However, they can be used to reproduce such data or, in other word, to conduct the process of electromagnetic infiltration. In the case of video signals (graphic mode of a computer, a laser printer), then the reproduced data can be presented in the form of graphic images that can be easily assimilated by people. Such images are transformed in order to find the data that is of interest. Reproduction of such data may lead to a disclosure of classified information.

There are many solutions that should counter the process of reproduction of such data. Such solutions are implemented in the design of equipment and influence, to a lesser or greater extent, the appearance of the equipment, as well as the related organizational methods. A new method that can be used in electromagnetic protection of process data is a software solution. It involves the use of appropriate computer fonts.

The article presents the possibilities related to shaping the form of video signals. For this purpose, appropriate shapes of the letter characters of computer fonts were recommended. Unlike characters in standard fonts (Arial and Times New Roman), they do not have the unique decorative elements (serif), such as hooks, connectors, heels, arches, and ribbons, and consist of only vertical and horizontal lines [5]. There are no slanted and crooked lines. Due to this, the characters are often very similar. This greatly contributes to the impossibility to differentiate between the letter characters in the reproduced image that is filled with noise and numerous disturbances. The graphic elements being searched, having the form of strings of letters, cannot be read. The digital image processing methods intended to improve the quality of the image are quite ineffective.

In search for graphic characters, such as computer font letter and digit characters, one can use methods based on the similarity of the standard with a portion of the analyzed image. However, when special fonts are used, the correlation method generates many false decisions, which also prevents reading text data.

Index Terms—Image analysis, image reconstruction, text recognition, data processing.

I. INTRODUCTION

When processing any data having the electronic form, radiated emission signals are generated. The nature of such emission can be compared to standard radio emission, e.g. broadcast by radio stations. However, when the main objective is not to radiate signals, such radiation is undesirable. The nature of such signals may correspond to the nature of the signals being processed [1]. In such situations, the emissions can be described as undesirable descriptions or byproducts. Quite often, such signals are referred to as compromising emanations. This is due to the fact that such signals have distinctive characteristics that correspond to the electrically processed data. Such characteristics include electric impulse duration, rising and falling edge duration, and impulse voltage amplitude value. The source of such signals can be any element of the device that processes data in electrical form. In particular, the graphic channel of a computer or the laser of a laser printer, can constitute such sources.

Due to the fact that such signals are not produced on purpose, it is not obvious that they can be received. This also applies to signals produced in the graphic channels of equipment. The suitability of such recorded signals is significant when the reception parameters correspond to the parameters of the video signal associated with the processed image. In the case of an image from the computer screen, this is the graphic mode of operation of the graphics card and the frequency of image refresh [2]. In the case of computer printers, it is the quality and speed of the print. When such parameters are taken into account in the receipt of undesirable signals, it may be possible to reproduce graphic data in the form of a standard image seen on a computer monitor or a printed sheet of paper [19].

Of huge importance to the data reproduction capability are the graphic form of the data and the characteristics of the Information Infiltration Channel (IIC). The IIC has the characteristics of a high-pass filter (differentiability):

$$x'(t) = \lim_{\Delta t \rightarrow 0} \frac{x(t + \Delta t) - x(t)}{\Delta t} \quad (1)$$

on the printer. The toner adheres to the drum only in locations where the laser beam struck the drum, which are positively charged by the laser. In the printing process, the paper, too, is positively charged which enables transferring the toner from the drum onto the surface of the paper.

Because of the interruptions of the laser beam, the beam strikes individual points which form lines in letter or digit characters. Each light emission is a source of a radiated emission signal which is used to reproduce the character using the compromising emanation signal. Note that the emission signal generated in this manner carries information not only on the vertical and slanted edges of printed characters [1], but also about the way they are filled (Fig. 2). This may increase the possibility to reproduce characters and to read the information. Further in this article we will demonstrate whether this applies to texts written with special (secure) fonts.

III. FORMS OF FONT CHARACTERS OBTAINED FROM UNDESIRABLE EMISSION SIGNALS.

The undesirable emission signals were recorded during printing of a text document that contained letter characters of standard fonts and of special fonts, in the arrangement presented in Figure 3. Signals were transmitted between the computer and the printer using the USB 2.0 standard and the quality of the document was 600 dpi x 600 dpi. The recorded signals had two frequencies: 700 MHz and 750 MHz.

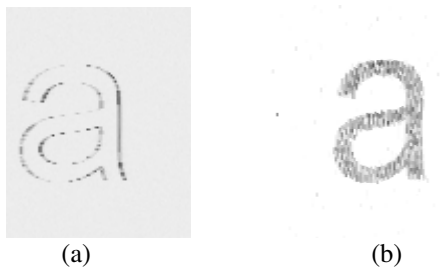


Fig. 2. An example form of a letter character obtained from a recorded undesirable emission signals whose source was a) a computer monitor and b) a laser printer.

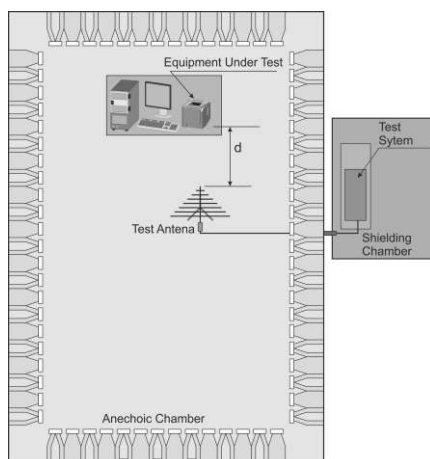


Fig. 3. The measuring system used to record the compromising emanation signals generated by the laser printer.

The special/secure computer fonts that protect processed text data from successful electromagnetic infiltration is the fonts whose lower case letters were shown, as an example, in Figure 4.

a) the special font 1



b) the special font 2



c) the special font 3

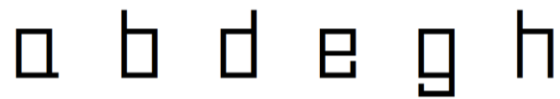
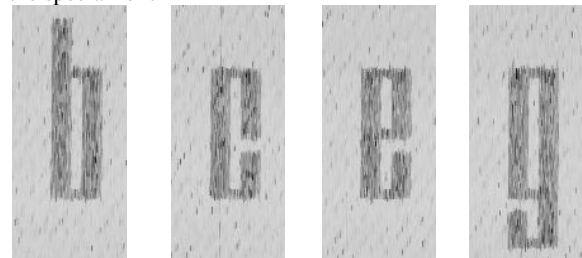


Fig. 4. The forms of letter characters of special fonts.

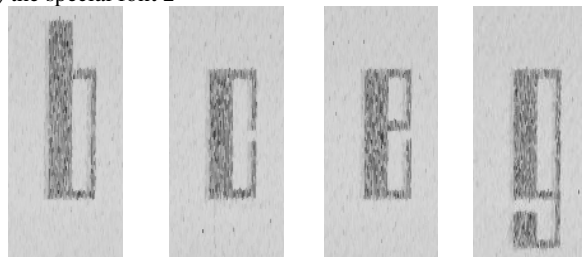
These are characters of one- (special font 3) and two-element (special fonts 1 and 2) typefaces. The letter characters have no unnecessary decorative elements (serif) that distinguish them from other letters in the same font.

The shapes of the graphic characters corresponding to letter characters of computer fonts, obtained in the process of image rastering from undesirable emission signals, are similar to their counterparts in the original text. Examples of such shapes are shown in Figure 5. However, those characters are very different from the graphic form of the characters reproduced from radiated emission signals whose source is the graphic channel of a VGA standard computer [6]. This is due to the nature of the video signal as the emission source, in the case of a computer monitor, and of the signal that controls the operation of the laser, in the case of a laser printer.

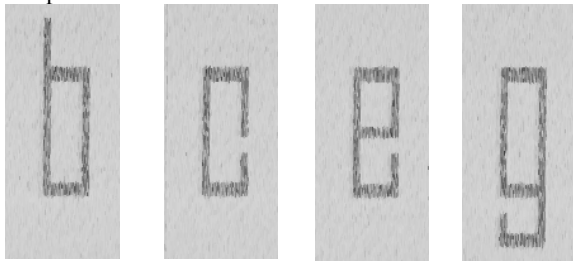
a) the special font 1



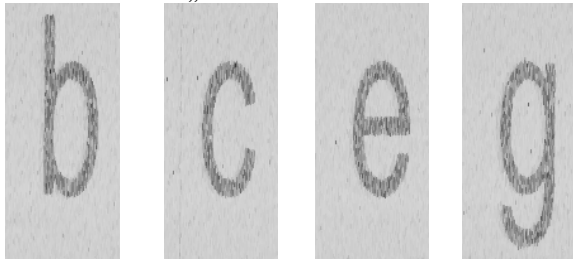
b) the special font 2



c) the special font 3



d) the traditional font „Arial”



e) the traditional font „Times New Roman”

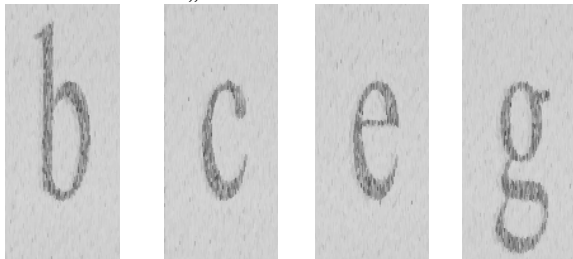
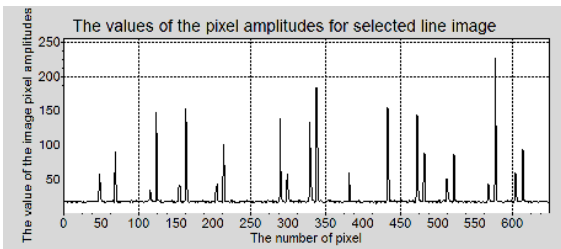


Fig. 5. Examples of letter characters of computer fonts reproduced from compromising emanation signals whose source was a laser printer - inverted images.

However, one must take note of the very important issue of the structure of the characters that are reproduced on the basis of recorded signals correlated with the signal of a printer laser.

a)



b)

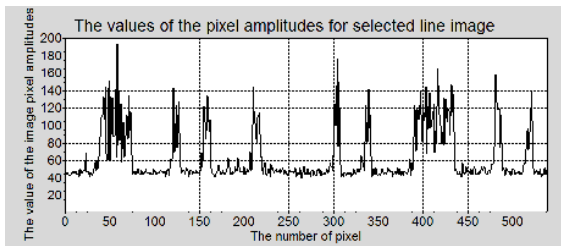
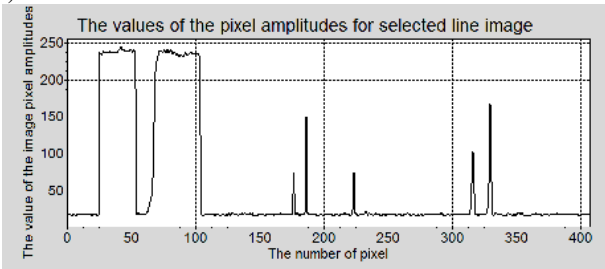


Fig. 6 Examples of paths of the compromising emanation signals that are recorded during operation (single horizontal lines (rows) of images) of a) the graphic channel of a computer and b) a laser printer, for the Arial font.

a)



b)

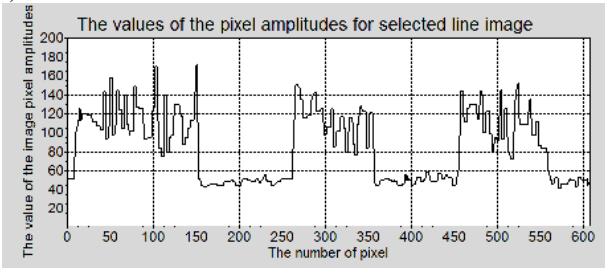


Fig. 7 Examples of paths of the compromising emanation signals that are recorded during operation (single vertical lines (columns) of images) of a) the graphic channel of a computer and b) a laser printer, for the Arial font.

As has been mentioned, each point forming a character, e.g. a letter, is preceded by an electric impulse that controls the operation of a printer laser. This influences the paths of the recorded undesirable emission signals (Fig. 6 and Fig. 7) and, consequently, the form of the reproduced characters.

The simple design of special font characters, which are limited to vertical and horizontal lines forming the letter, the lack of oblique and curved lines, makes their visual analysis very difficult. Noise in the images, interferences, and low levels of recorded signals make such analysis even more complicated.

IV. REAL IMAGES CONTAINING LETTER CHARACTERS

The analyses were based on signals on the output of the IIC, which were real signals. The results obtained in each case demonstrated that by properly selecting the shapes of the letter characters one can influence the degree of success in reproduction of data carried in undesirable emissions. In the case of the recommended shapes of letter characters, reproduction of characters is connected with the degree of their mutual similarity.

Figure 8 shows the real image and the image obtained using the rastering method from a recorded emission signal correlated with the operation of the laser of a laser printer. The printed document contained the same text written using the five analyzed computer fonts. The obtained image is not very legible and its quality is very poor. In such cases, the process of signal classification uses digital image processing methods whose operating algorithms are modified for the purpose of electromagnetic infiltration. The methods used with reference to the images obtained from signals whose source was the printer of a laser were median filtration of mirror size 3x3 and thresholding of the value of the

amplitude of the pixels forming the reproduced image, with the threshold values equal to 40 (lower threshold) and 218 (upper threshold). The results of the operations are shown in Figures 9 and 10.

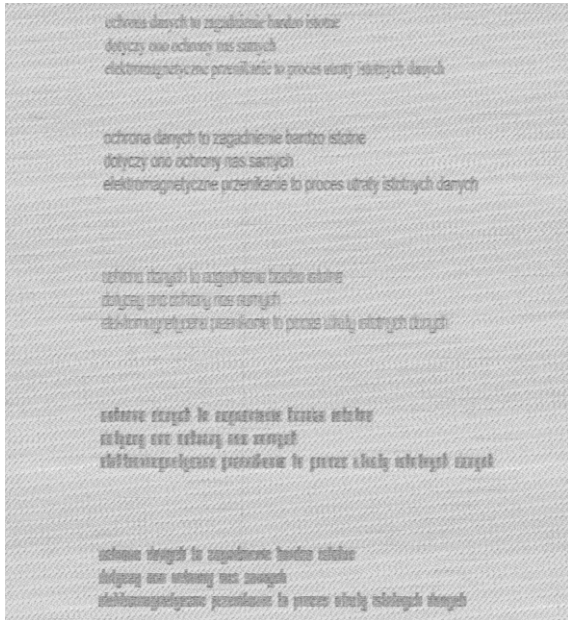


Fig. 8. A part of the real image obtained from a compromising emission signal generated by a laser printer without implementing digital processing methods aimed to improve the quality of the image, containing text written with the five analyzed computer fonts - inverted image.

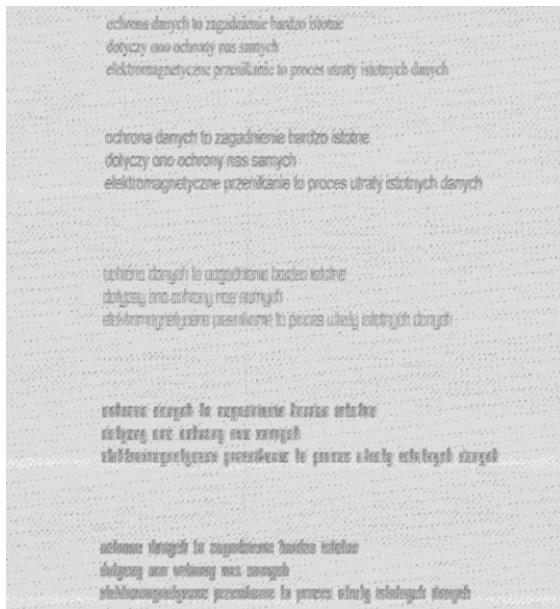


Fig. 9. A part of the real image obtained from a compromising emission signal generated by a laser printer after two operations of median filtration with the window dimensions of 3x3, containing text written with the five analyzed computer fonts - inverted image.

Because each image contains five texts, the digital image processing method used affect each of the characters contained in the image in the same manner. The methods used enabled obtaining images of better contrast and legibility. Nevertheless, some character strings are not legible.

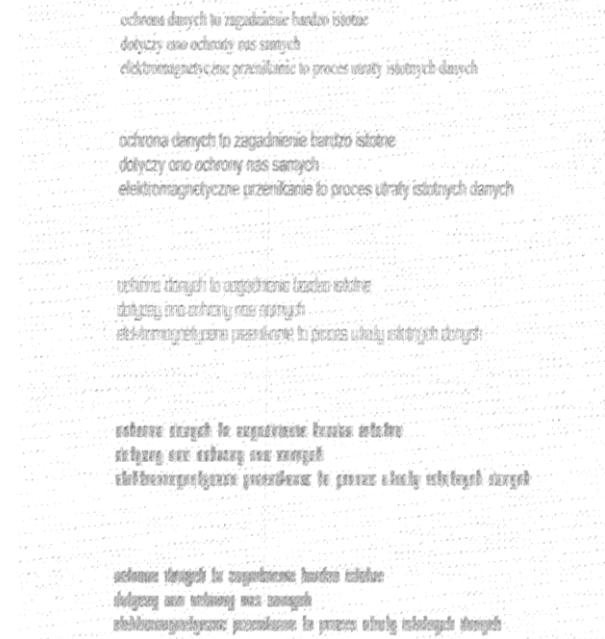


Fig. 10. A part of the real image obtained from a compromising emission signal generated by a laser printer after two operations of median filtration with the window dimensions of 3x3 and thresholding of the amplitude values with an upper and lower threshold, containing text written with the five analyzed computer fonts - inverted image

In the images, one can clearly see that the standard font characters are legible and, in most cases, can be distinguished from others. On the other hand, the special fonts shown in Figure 8 are not legible. The use of digital image process methods makes attempts at reading the text written using special font no. 3 possible. Nevertheless, this may prove to be difficult. The other special fonts (1 and 2) make the text contained in the image completely unsuitable for reading using the visual method. In particular, special font 3 is a string of dark spots that form the outlines of the rows of the text. The use of single line spacing between the rows would make the image completely blurry.

Nevertheless, visual analysis of images of poor quality is often insufficient. One must search and verify the results using supporting methods. Most often such methods involve numerical analyses supported decision-wise with appropriate conditions and criteria. One of the conditions used in the process of searching for graphic characters in an image is the threshold value of the mutual correlation coefficient. Based on this value, one can determine the value of the CER for each computer font.

V. CHARACTER CORRELATION COEFFICIENT OF COMPUTER FONTS AND CHARACTER ERROR RATE

Due to strong disturbances and noise in images, their visual analysis prevents indicating the special font characters present in images. However, one may attempt to support the relevant decisions by using appropriate conditions and criteria associated with computer analysis.

As has been mentioned before, a threshold value of the character correlation coefficient has been used, for which

it is assumed that at least 90% of specific characters will be recognized. This enables determination of the CER value of the analyzed fonts with regards to the number of incorrect decisions associated with recognition of a given character. The values of the character correlation coefficient are determined using the following formula:

$$R^z_{j,i} = \frac{\sum_{n=0}^{N_w-1} \sum_{m=0}^{M_w-1} A \cdot B}{\sqrt{\sum_{n=0}^{N_w-1} \sum_{m=0}^{M_w-1} A^2 \cdot \sum_{n=0}^{N_w-1} \sum_{m=0}^{M_w-1} B^2}} \quad (2)$$

where:

$$\bar{x}_{j,i} = \frac{1}{N_w \cdot M_w} \sum_{n=0}^{N_w-1} \sum_{m=0}^{M_w-1} C \quad (3)$$

$$\bar{y} = \frac{1}{N_w \cdot M_w} \sum_{n=0}^{N_w-1} \sum_{m=0}^{M_w-1} y_{n,m} \quad (4)$$

$$A = C - \bar{x}_{j,i} \quad (5)$$

$$B = y_{n,m} - \bar{y} \quad (6)$$

$$C = x_{n+j \cdot N_w, m+i \cdot N_w} \quad (7)$$

and $0 \leq j \leq b-1$, $0 \leq i \leq d-1$, $b = N_a / N_w$ – number of rows in the correlation image, $d = M_a / M_w$ – number of columns in the correlation image, $x_{n,m}$ – amplitudes of pixels of the analyzed image containing alphabetic characters, $y_{n,m}$ – amplitudes of pixels of the reference image (a single alphabetic character), M_a – number of columns in the analyzed image containing alphabetic characters, N_a – number of rows in the analyzed image containing alphabetic characters, M_w – number of columns in the reference image (a single alphabetic character), N_w – number of rows in the reference image (a single alphabetic character), i – number of the correlation image column, m – number of the reference image column (a single alphabetic character), j – number of the correlation image row, n – number of the reference image row (a single alphabetic character), as well as the algorithm shown in Figure 11.

The obtained values of the character correlation coefficient for the individual characters within a given font are shown in Tables 1 and 2.

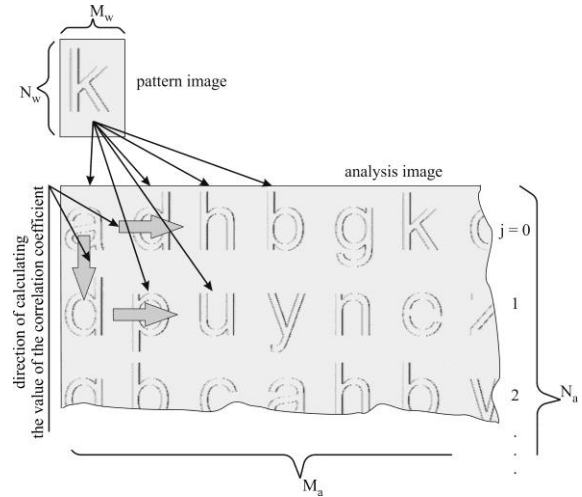


Fig. 11. The rule for calculating the character correlation coefficient R^z for computer font letter characters (in this case, letter “k” shown in the figure) in order to determine the CER.

Table 1. Number of characters (lower case letters) for special fonts whose values of the R^z coefficient of correlation between the lower case letter characters at the output of the IIC are within the defined intervals.

Value of the correlation coefficient	Number of lower case letter characters of a given font for which the value of the R^z correlation coefficient (similarity to another character of the same font) is within a defined interval.		
	Special font 1	Special font 2	Special font 3
$R < 0.2$	0	0	0
$0.2 \leq R < 0.4$	5	0	16
$0.4 \leq R < 0.7$	135	128	142
$0.7 \leq R < 0.9$	58	70	40
$R \geq 0.9$	0	0	0

Table 2. Number of characters (lower case letters) for standard fonts whose values of the RZ coefficient of correlation between the lower case letter characters at the output of the IIC are within the defined intervals.

Value of the correlation coefficient	Number of lower case letter characters of a given font for which the value of the R^z correlation coefficient (similarity to another character of the same font) is within a defined interval.	
	Arial	Times New Roman
$R < 0.2$	8	4
$0.2 \leq R < 0.4$	80	89
$0.4 \leq R < 0.7$	109	103
$0.7 \leq R < 0.9$	1	2
$R \geq 0.9$	0	0

Because of the noise in the images, the character correlation coefficient for special fonts may have larger values for non-conforming characters (the standard character does not conform with the character contained in the part of the analyzed image) than for conforming characters (table 3). As a result, this may lead to incorrect recognition of characters based on the condition that the correlation coefficient must have an appropriate value.

Table 3. The maximum values of the character correlation coefficient for conforming and non-conforming characters for special fonts, $SNR = -6dB$.

Character	Special font 1		Special font 2	
	$R_{b,max}^Z$	$R_{l,max}^Z$	$R_{b,max}^Z$	$R_{l,max}^Z$
a	0.2105	0.1896	0.1750	0.1855
b	0.2012	0.1917	0.1955	0.1987
c	0.1962	0.1876	0.1896	0.1887
d	0.2001	0.1852	0.1916	0.1836
e	0.1948	0.1996	0.1787	0.1852
g	0.1938	0.1895	0.1951	0.1893
h	0.2025	0.2031	0.1872	0.2028
i	0.1783	0.1406	0.1861	0.1692
k	0.1982	0.1926	0.1872	0.1992
l	0.1796	0.1478	0.1686	0.1850
n	0.1751	0.1927	0.1903	0.1951
o	0.1891	0.1895	0.1880	0.1913
p	0.1990	0.1742	0.1865	0.1873
r	0.1846	0.1545	0.1779	0.1671
s	0.1884	0.1735	0.1870	0.1832
t	0.1645	0.1624	0.1786	0.1801
u	0.1922	0.1920	0.1915	0.1874
y	0.1852	0.1917	0.1905	0.2008
z	0.1857	0.1803	0.1872	0.1911

Thus, assuming a certain value of the R^Z character correlation coefficient above which the decision is made to recognize a character, one may conduct the process of search for letter or digit characters in the reproduced image. The problem that remains is adoption of appropriate threshold value of the R_{thr}^Z parameter. In the analyses that were performed, it was assumed that the percentage of the recognized characters should not be less than 90% of all the characters being searched. The relevant calculations of the R_{thr}^Z and the CER were conducted on images that contained letter characters of the different fonts with the equal number of each character (Fig. 12), for which the signal to noise ratio was equal to $-6dB$. The locations of characters in the image were random. The CER value was determined using the following formula:

$$CER = \frac{c+k}{f} = \frac{c+(r-e)}{f} \quad (8)$$

where: r – the number of a given character present in the analyzed image (for the analyzed cases $r = 31$), c – the number of characters recognized incorrectly, e – the number of characters recognized correctly, k – the

number of characters that were not recognized but were searched ($k = r - e$), f – the number of all characters present in an image (for the analyzed cases $f = 660$). In the case of the analyzed images and the assumed values of the r and f parameters, formula (8) has the following form:

$$CER = \frac{c+(31-e)}{660} \quad (9)$$

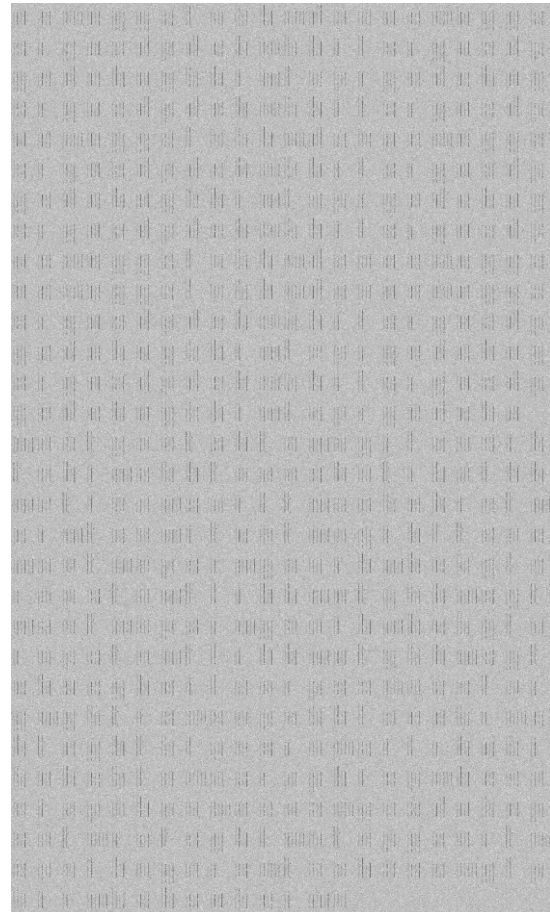


Fig. 12. An example of an analyzed image containing letter characters of special font 1, $SNR = -6dB$.

The results of the analyses demonstrate that the degree of similarity between characters of special fonts is much higher than that between characters of standard fonts. This greatly affects the possibility to identify characters in images.

A high degree of similarity between characters translates into a large number of errors defined by the CER . In the case of special fonts, the values are $300 \cdot 10^{-3}$ or $400 \cdot 10^{-3}$ (table 4) and in the case of standard fonts – only $60 \cdot 10^{-3}$ (table 5).

VI. CONCLUSION

The article presents the shapes of special font characters and their influence of the forms of graphic characters reproduced from undesirable emission signals.

The source of signal emission was a laser printer and the signal was recorded in the course of printing of a text document.

Table 4. CER values for special fonts, for $SNR = -6dB$

Character	Special font 1	Special font 2	Special font 3
a	280	589	221
b	127	115	182
c	484	654	170
d	68	409	163
e	433	715	314
g	46	105	61
h	92	221	131
i	10	430	19
k	284	211	34
l	5	278	151
n	426	443	430
o	311	693	577
p	299	178	341
r	3	615	370
s	267	640	166
t	139	161	24
u	387	438	316
y	66	76	54
z	358	418	267

Table 5. CER values for special fonts, for $SNR = -6dB$

Character	Arial	Times New Roman
a	5	7
b	0	14
c	73	24
d	0	0
e	7	14
g	5	0
h	3	34
i	53	46
k	5	5
l	143	8
n	49	87
o	32	51
p	3	5
r	34	27
s	2	2
t	7	12
u	5	14
y	0	2
z	5	5

The reproduced special font characters were compared with standard font characters (Arial and Times New Roman). The results of the analyses demonstrate the presence of a strict relationship between the form of a letter character and the possibility of its reproduction by rastering from recorded undesirable emission signals. Traditional font letter characters have many graphic elements that make them different from one another.

Those elements are serifs, heels, rounded elements, arches, etc. Special font characters, being as simple as possible, contain no such elements. Also, they consist only of vertical and horizontal lines. Such details are of great importance to the degree of similarities between letter characters. The high level of correlation between such characters makes it impossible to distinguish between different characters on the receiving side. This is demonstrated both in a visual analysis and in a numerical analysis based on determination of R_{thr}^Z threshold values of the character correlation coefficient that require recognition of at least 90% of all the characters being searched R^Z values for the reference character and the part of the analyzed image that are greater than the R_{thr}^Z value determine recognition of the character being searched. However, in particular in the case of special fonts, this does not always indicate conformity of the reference character to the character contained in the analyzed portion of the image. Thus, the decision is incorrect. The indicated character is not present in the specific place in the image. Consequently, the value of the Character Error Rate parameter was calculated. The parameter indicates the degree to which special fonts are more resistant to recognition using the correlation method or, in other words, how many incorrect decisions are made. The data obtained demonstrates much higher resistance of special fonts to correct recognition of characters. The CER values for those fonts were equal to approximately $400 \cdot 10^{-3}$, compared to $60 \cdot 10^{-3}$ in the case of standard fonts. This phenomenon can be used when designing computer font characters to be used in electromagnetic protection of classified information. The first such fonts are indicated in the article. This problem is particularly important given the common use of computers for processing classified information in having the form of text documents.

REFERENCES

- [1] Kubiak I., "Possibilities of reconstruct of data text from unwanted signal emissions using character correlation method – standard DVI", Telecommunication Review and Telecommunication News, 2014-2-3.
- [2] Kubiak I., "Digital processing methods of images and signals in electromagnetic infiltration process", Image Processing and Communications, vol. 18, no. 1, pp. 5-16, 2013, DOI: 10.2478/v10248-012-0070-7, ISSN: 1425-140X.
- [3] Kubiak I., Grzesiak K., Musial S., Przybysz A., „The generator of half-tone screening for the process of electromagnetic infiltration”, Military University of Technology 2012, ISBN 978-83-62954-28-5.
- [4] Kubiak I., "Identification of waveforms for electromagnetic fields on the basis of its measurements within a confined bandwidth of reception", PhD Thesis 2000.
- [5] Kubiak I., "Analysis methods and digital processing images in electromagnetic infiltration process", Military University of Technology 2013, ISBN 978-83-62954-86-5.
- [6] Richard G. Lyons: Understanding Digital Signal Processing (2nd Edition), 2004.
- [7] Zielinski T.: "Digital Signal Processing",

- Communication and Transport Publisher, 2009.
- [8] Deok J. Park, Kwon M. Nam: "Multiresolution Edge Detection Techniques, Pattern Recognition", Vol. 28, 1995.
- [9] Pratt K.: "Digital Image Processing", Wiley& Sons, New York 1978.
- [10] Michael J. McCarthy, "The Pentagon worries that spies can see its computer screens, someone could watch what's on your VDT", The Wall Street Journal, 07.08.2000.
- [11] Allen R. L., Mills D., "Signal Analysis: Time, Frequency, Scale, and Structure", 2004.
- [12] Hong Zeng, Dual image processing algorithms and parameter optimization, Seventh International Conference on Natural Computation (ICNC), Shanghai 2011, Conference materials volume 2, p.946-950, ISSN 2157-9555.
- [13] Sohi D.S., Application to enhance the teaching and understanding of basic image processing techniques, Southeastcon 2000, Naswille, p. 413-416, ISBN 0-7803-6312-4.
- [14] Mitra S.K., Image processing using quadratic volterra filters, 5th International Conference on Computers and Devices for Communication (CODEC), Kolkata 2012, ISBN 978-1-4673-2619-3.
- [15] Grzesiak K., Przybysz A., "Emission security of laser printers", MCC 2010: Military Communications and Information Systems Conference, Wroclaw 2010.



Ireneusz Kubiak was born in Poland. Now he is working in Military Communication Institute (MCI) in Żegrze in Poland. He is head of Electromagnetic Compatibility Division (since 2004) and EMC Laboratory of the MCI (since 2004). In 1992 he graduated the Military University of Technology (faculty of electronics), in 1993 - Warsaw University (faculty of Mathematics, Mechanics and Computer Science). In 2000 he got a degree of doctor of technical sciences. Currently he is engaged in research and analysis of electromagnetic emissions, in particular, correlated with classified information. In addition, a lot of time he dedicates to electromagnetic compatibility. He is the author of four books (2009, 2012, 2013, 2014) which describes aspects of protection of information against electromagnetic penetration.

How to cite this paper: Ireneusz Kubiak, "The Unwanted Emission Signals in the Context of the Reconstruct Possibility of Data Graphics", IJIGSP, vol.6, no.11, pp.1-9, 2014. DOI: 10.5815/ijigsp.2014.11.01