Ireneusz KUBIAK

Wojskowy Instytut Łączności

Font komputerowy odporny na proces infiltracji elektromagnetycznej

Streszczenie. *Ochrona przed elektromagnetycznym przenikaniem informacji to nie tylko rozwiązania konstrukcyjne urządzeń czy też strefy ochrony fizycznej. Stosowane rozwiązania konstrukcyjne wpływają na masę urządzeń i niedogodności ergonomiczne. Poszukuje się więc inne technologie wspomagające ochronę elektromagnetyczną informacji. Jednym z nowych rozwiązań są fonty komputerowe o specjalnych kształtach. Właściwości Kanału Przenikania Informacji powodują, że tekst pisany fontem bezpiecznym jest odporny na proces infiltracji elektromagnetycznej.*

Abstract. *Protection of information against electromagnetic leakage it is not only construction solutions of devices or physical protection zones. The use of these construction solutions influence equipment mass and ergonomic inconveniences. Other technologies supporting protection of information are looked for. One of new solutions are special shapes of computer fonts. The characteristics of Leakage Information Channel causes that text written special computer font is resistant to electromagnetic infiltration process.* ***(The computer font resistant to electromagnetic infiltration process)***

Słowa kluczowe: font bezpieczny elektromagnetycznie, emisja ujawniająca, proces infiltracji elektromagnetycznej, ochrona informacji.
Keywords: electromagnetic safety font, compromising emanation, electromagnetic infiltration process, protection of information

doi:10.12915/pe.2014.06.40

Wstęp

Ochrona przed elektromagnetycznym przenikaniem informacji wymaga szereg przedsięwzięć organizacyjnych, jak i konstrukcyjnych. Obecnie wykorzystywane metody inżynierii kompatybilności elektromagnetycznej są na bardzo wysokim poziomie zaawansowania techno-logicznego. Ich implementacja w różnego typu urządzeniach nie wpływa na ich wygląd zewnętrzny, ale zwiększa ich masę i czasami wprowadza ograniczenia ergonomiczne. Ponadto, stosowane zabezpieczenia powodują wzrost kosztów zakupu sprzętu informatycznego przeznaczonego do przetwarzania informacji niejawnych. Szuka się zatem innych rozwiązań chroniących przetwarzane dane. Rozwiązaniami tymi mogą być np. przedsięwzięcia softwarowe polegające na stosowaniu specjalnego fontu komputerowego. Kształt znaków liter i cyfr takiego fontu musi uwzględniać charakter Kanału Przenikania Informacji (KPI). Wówczas odtworzone z sygnału emisji ujawniającej dane (obrazy) będą nieczytelne, przez co przetwarzane informacje będą bezpieczne elektromagnetycznie. Nawet stosowanie metod cyfrowego przetwarzania sygnałów i obrazów (filtracja logiczna, filtracja medianowa, progowanie wartości amplitud pikseli), nie będzie przynosić oczekiwanych efektów, którymi jest poprawa jakości odtwarzanych obrazów umożliwiająca odczyt danych zwłaszcza danych tekstowych.

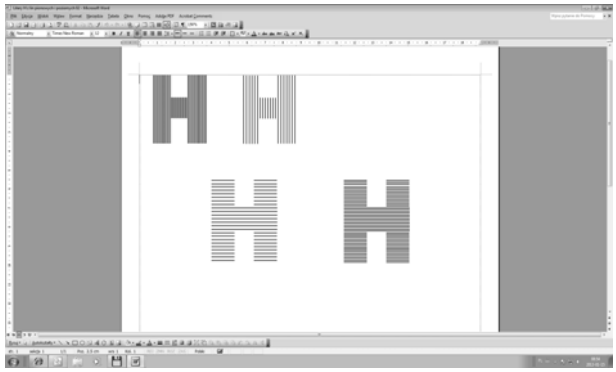
Kanał Przenikania Informacji a postać sygnału emisji ujawniającej

Rozpatrując zagadnienia związane z możliwością prowadzenia procesu infiltracji elektromagnetycznej, wykorzystując do tego celu wyszukane metody cyfrowego przetwarzania sygnałów i obrazów, nie można zapominać o istotnym czynniku wpływającym na ten proces, jakim jest KPI. Poza elementami technicznymi stosowanymi w rozwiązaniach konstrukcyjnych urządzeń (ograniczenie poziomów emisji niepożądanych), KPI wraz z torem odbiorczym miernika pomiarowego wprowadza zniekształcenia tych sygnałów. Zniekształcenia te spowodowane są charakterystyczną dla tego typu kanałów cechą, jaką jest ich różniczkowalność, definiowaną następująco:

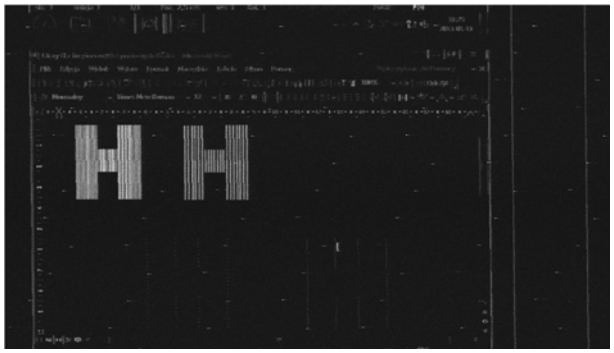
$$x'(t) = \lim_{\Delta t \to 0} \frac{x(t+\Delta t) - x(t)}{\Delta t}, \tag{1}$$

gdzie: $x'(t)$ – sygnał na wyjściu KPI, $x(t)$ – sygnał emisji ujawniającej (sygnał na wejściu KPI), a także szumami i zaburzeniami pochodzenia naturalnego, jak i wynikające z działalności człowieka ($s(t)$).

a) pierwotne znaki wyświetlane no monitorze komputera



KPI



Rys.1. Wpływ KPI na postać obrazu

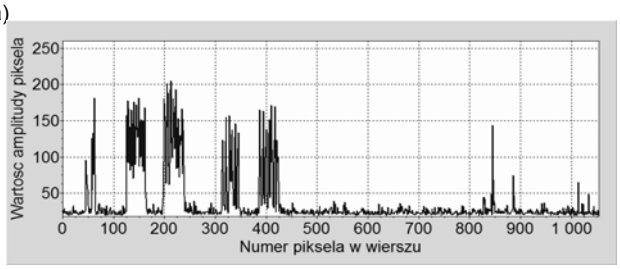
W praktyce, postać sygnału jaką dysponujemy i która podlega następnie cyfrowemu przetwarzaniu, znacznie różni się od postaci sygnału $x'(t)$. Spowodowane jest to dodatkowo zniekształceniem jakie wprowadza odbiornik pomiarowy. Na jego wyjściu otrzymujemy sygnał x"(t) opisany zależnością:

$$x''(t) = \left| \lim_{\Delta t \to 0} \frac{x(t+\Delta t) - x(t)}{\Delta t} + s(t) \right|. \tag{2}$$

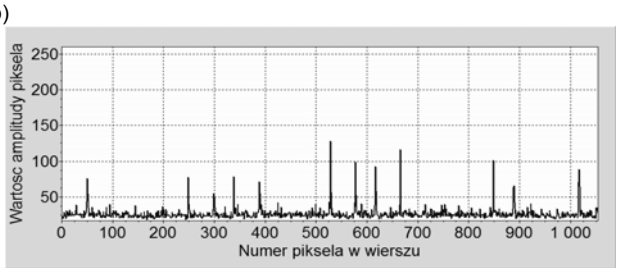
KPI zachowując się jak filtr górnoprzepustowy tłumi składowe niskich częstotliwości, przepuszczając jednocześnie składowe sygnału o wyższych częstotliwościach. O kształcie przebiegu czasowego sygnału decydują więc składowe sygnału o wyższych częstotliwościach. Jest to szczególnie istotna właściwość, z punktu widzenia projektowania kształtów znaków fontów wspomagających ochronę przed elektromagnetycznym przenikaniem.

Zwróćmy uwagę w jaki sposób KPI wpływa na postać odtwarzanych znaków graficznych z rejestrowanych sygnałów emisji ujawniających. Analizując rzeczywiste obrazy (rys.1) zauważamy potwierdzenie omawianych wcześniej właściwości różniczkujących KPI. Litery „H" budowane z linii poziomych nie są zauważane na odtwarzanym obrazie. Można jedynie spostrzec zarysy początków i końców tych linii.

a)



b)



Rys.2. Przebieg zmienności wartości amplitud pikseli budujących obraz po KPI (rys.1b) a) zawierający znaki liter „H" budowanych z linii pionowych, b) zawierający znaki liter „H" budowanych z linii poziomych

Jednak i one w wielu przypadkach, w towarzystwie występujących zaburzeń, są maskowane przez co i niezauważalne (rys.1). Linie pionowe o szerokości kilku pikseli, jako sygnał szybkozmienny, ulega nieznacznym zniekształceniom przez co informacja niesiona przez sygnał dociera do odbiornika.

Postać bezpiecznego fontu komputerowego

Wymagania na font bezpieczny

Najłatwiejszym i być może najtańszym rozwiązaniem ochrony informacji przed elektromagnetycznym przenikaniem może okazać się stosowanie rozwiązań softwarowych. Nie należy jednak przez to rozumieć rozwiązań kryptograficznych. Tego typu podejście do przetwarzania danych graficznych, które muszą być wyświetlane w sposób jawny na monitorze komputera, nie może być stosowane. Rozwiązania softwarowe to przede wszystkim rozwiązania dotyczące dokumentów tekstowych, mające na celu stosowanie fontów komputerowych o odpowiednim kroju. Fonty takie dla zapewnienia wymagań w zakresie ochrony przed elektromagnetycznym przenikaniem informacji oraz pozytywnej opinii potencjalnych użytkowników, muszą spełniać odpowiednie założenia. Do założeń tych zaliczamy:

a) odpowiednia czytelność

Kształty znaków muszą być pozbawione elementów upiększających, które wpływają jednocześnie na charakter każdego znaku. Mimo tego, każdy znak fontu powinien być czytelny i rozróżnialny. Podobieństwo między znakami powinno być odpowiednio duże ale jednocześnie, aby nie wywoływało u czytelnika trudności z interpretacją znaku.

b) wyrazistość i kontrast

Najgorszym rozwiązaniem dla czytelnika jest brak kontrastu czyli brak wyrazistości znaków graficznych. Jednym z rozwiązań, które proponowano już kilka lat temu było wydłużenie czasów narastania impulsów występujących w sygnale wideo i decydujących o wyświetlanej grafice na monitorze. Fonty bezpieczne muszą zapewniać typową wyrazistość i kontrast znaków spotykane w fontach tradycyjnych.

c) ograniczenia skuteczności odtworzeniowych

Stosowanie typowego fontu np. „Arial" lub „Times New Roman" pokazuje, że w przypadku wystąpienia emisji ujawniających pozwalających na identyfikowanie informacji niejawnych, możliwe jest jej odtworzenie poprzez wykorzystanie metod cyfrowego przetwarzania obrazów. Nowy font powinien być pozbawiony tej właściwości.

d) brak skuteczności oprogramowania OCR (ang. Optical Character Recognition)

Specjalny font może spełniać dwojakie zadanie. Po pierwsze, po przejściu przez KPI jego cechy dystynktywne mogą zostać gubione, przez co jego odtworzenie może być niemożliwe lub z bardzo dużym prawdopodobieństwem niepewności. Po drugie – zakładane podobieństwo znaków liter musi uniemożliwiać przenoszenie wersji papierowej dokumentu na wersję elektroniczną za pomocą dostępnych oprogramowań typu OCR.

e) ograniczenie wykorzystania korelacji dwuwymiarowej

Bezpieczeństwo elektromagnetyczne fontu komputerowego to również jego odporność na rozpoznawanie znaków metodą korelacji dwuwymiarowej po przejściu sygnału emisji ujawniającej przez KPI. W takim przypadku należy zapewnić aby po stronie odbiorczej znaki posiadały jak najwięcej wspólnych cech upodobniających poszczególne znaki do siebie. Można to osiągnąć poprzez odpowiednie podobieństwo znaków pierwotnych.

Kształty znaków fontu bezpiecznego

Uwzględniając właściwości KPI oraz przyjęte założenia dotyczące funkcji jakie powinien spełniać font w wykonaniu specjalnym, zaprojektowano dwa rodzaje fontów bezpiecznych. Obrazy fontów różnią się między sobą, zapewniając jednocześnie odpowiedni stopień ochrony elektromagnetycznej.

Font bezpieczny to przede wszystkim znaki proste, maksymalnie podobne do siebie. Brak elementów budowy jednego znaku powoduje, że może być on rozpoznawany jako inny znak litery. Na rysunku 3 przedstawiono przykładowe znaki liter fontów specjalnych.

a) font „Bezpieczny Symetryczny"

b) font „Bezpieczny Niesymetryczny"

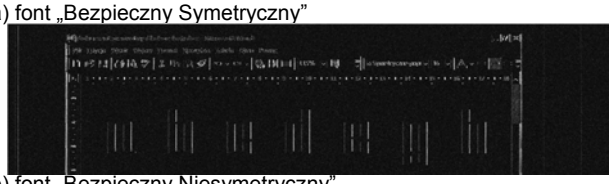
Rys.3. Postacie fontów bezpiecznych

Wpływ KPI na kształt znaków fontu bezpiecznego

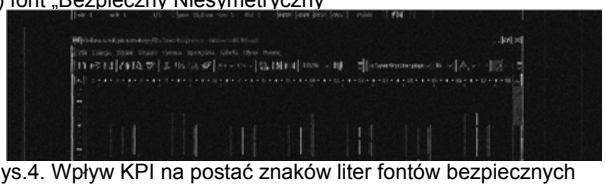
Postacie znaków przedstawionych na rysunku 3 poddano działaniu KPI, w wyniku czego otrzymano sygnały emisji ujawniającej, z których odtworzono, metodą rastrowania,

obraz pierwotny (rys.4). Dla porównania czytelności odtwarzanych znaków liter, na rysunku 5 zamieszczono odtworzone obrazy, które odpowiadają obrazom pierwotnym zawierającym znaki fontów „Arial” i „Times New Roman”. Jak możemy zauważyć, czytelność znaków fontów tradycyjnych jest dużo większa niż czytelność znaków fontów bezpiecznych.

a) font „Bezpieczny Symetryczny”

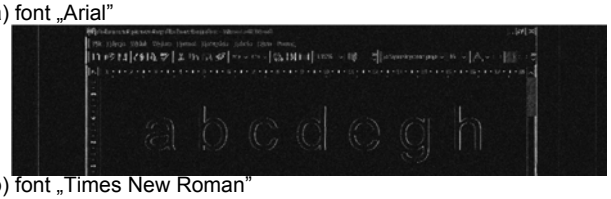


b) font „Bezpieczny Niesymetryczny”



Rys.4. Wpływ KPI na postać znaków liter fontów bezpiecznych

a) font „Arial”



b) font „Times New Roman”

Rys.5. Wpływ KPI na postać znaków liter fontów tradycyjnych

Próby odtworzenia znaków liter fontu bezpiecznego

Korelacja między znakami

Do określenia stopnia podobieństwa między znakami w ramach poszczególnych fontów posłużono się macierzą znaków, której fragment przedstawiono na rysunku 6. Współczynnik korelacji znakowej R^Z liczony jest dla odpowiednich fragmentów obrazu analizowanego, zawierających poszukiwane znaki, a obrazem wzorca zawierającym poszukiwany znak, zgodnie z zależnością:

$$R^z_{j,i} = \frac{\sum_{n=0}^{N_w-1}\sum_{m=0}^{M_w-1} K\cdot\left(y_{n,m}-\bar{y}\right)}{\sqrt{\sum_{n=0}^{N_w-1}\sum_{m=0}^{M_w-1} K^2\cdot\sum_{n=0}^{N_w-1}\sum_{m=0}^{M_w-1}\left(y_{n,m}-\bar{y}\right)^2}} \tag{3}$$

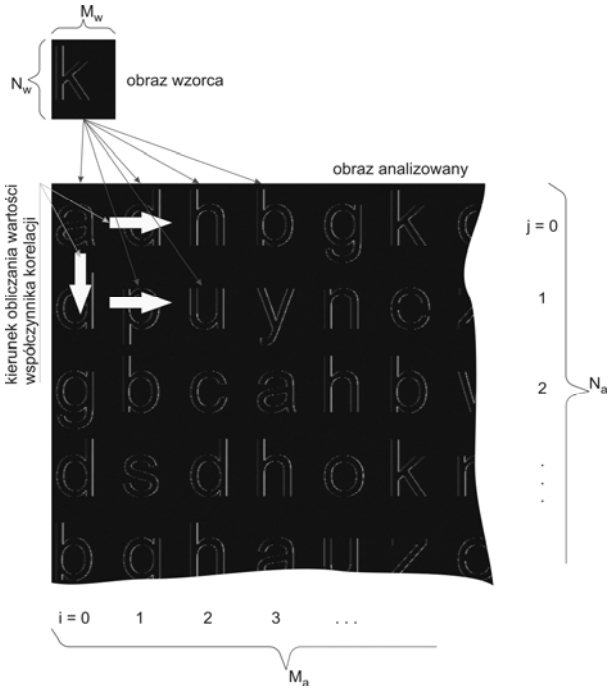
gdzie:

$$K = \left(x_{n+j\cdot N_w,m+i\cdot N_w}-\bar{x}_{j,i}\right), \tag{4}$$

$$\bar{x}_{j,i} = \frac{1}{N_w\cdot M_w}\sum_{n=0}^{N_w-1}\sum_{m=0}^{M_w-1} x_{n+j\cdot N_w,m+i\cdot M_w}\,, \tag{5}$$

$$\bar{y} = \frac{1}{N_w\cdot M_w}\sum_{n=0}^{N_w-1}\sum_{m=0}^{M_w-1} y_{n,m}\,, \tag{6}$$

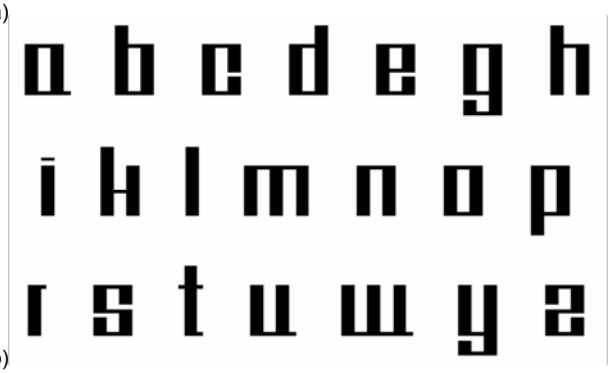
oraz $0 \le j \le b-1$, $0 \le i \le d-1$, $b = N_a/N_w$, $d = M_a/M_w$, $x_{n,m}$ – amplitudy pikseli obrazu analizowanego, $y_{n,m}$ – amplitudy pikseli obrazu wzorcowego, M_a – liczba kolumn obrazu analizowanego; N_a – liczba wierszy obrazu analizowanego, M_w – liczba kolumn obrazu wzorcowego, N_w – liczba wierszy obrazu wzorcowego, i – numer kolumny tekstu obrazu korelacji, m – numer kolumny obrazu wzorcowego, j – numer wiersza tekstu obrazu korelacji, n – numer wiersza obrazu wzorcowego.



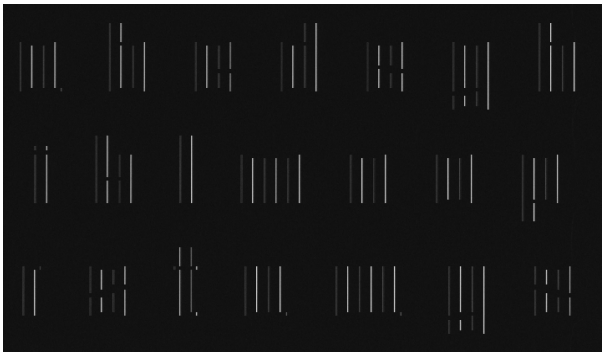
Rys.6. Reguła obliczania współczynnika korelacji znakowej R^Z dla znaków liter fontów komputerowych

Celem wyznaczenia wartości współczynnika korelacji $R^Z_{j,i}$ wykorzystano obrazy, których przykład zamieszczono na rysunku 7.

a)



b)



Rys.7. Obraz pierwotny (a) i odtworzony (b) z sygnału emisji ujawniającej zawierające znaki liter fontu „Bezpieczny Symetryczny”

Wyniki jakie otrzymano zawarto w tabeli 1 oraz na rysunku 8. Dane tablicowe mówią o liczbie znaków dla

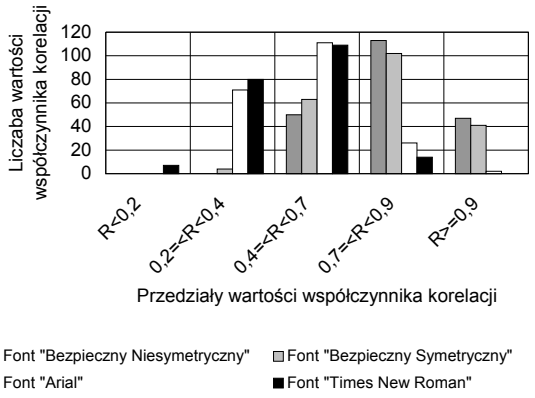
danego fontu, których wartość współczynnika korelacji R^Z mieści się w zadanych przedziałach. Z przytoczonych danych widać, że założenie dotyczące maksymalizacji podobieństwa znaków fontów bezpiecznych zostało spełnione. Dla tego typu fontów wartość R^Z to przede wszystkim $[0{,}7;0{,}9)$. Fonty tradycyjne to z kolei $[0{,}4;0{,}7)$.

Tabela 1. Liczba znaków fontu, dla którego wartość współczynnika korelacji na wejściu KPI zawiera się w zadanym przedziale wartości R^Z

Wartość współczynnika korelacji	Liczba znaków danego fontu, dla których wartość współczynnika korelacji R^Z (podobieństwa do innego znaku danego fontu) mieści się w zadanym przedziale			
	BN	BS	Arial	Times New Roman
$R < 0.2$	0	0	0	7
$0.2 \le R < 0.4$	0	4	71	80
$0.4 \le R < 0.7$	50	63	111	109
$0.7 \le R < 0.9$	113	102	26	14
$R \ge 0.9$	47	41	2	0

BN – font „Bezpieczny Niesymetryczny", BS – font „Bezpieczny Symetryczny"

Nasuwa się więc pytanie: na ile uzyskany stopień podobieństwa znaków wpływa na korelację między znakami na wyjściu KPI?



Rys.8. Graficzne przedstawienie liczba znaków danego fontu, dla którego wartość współczynnika korelacji na wejściu KPI zawiera się w zadanym przedziale wartości R^Z

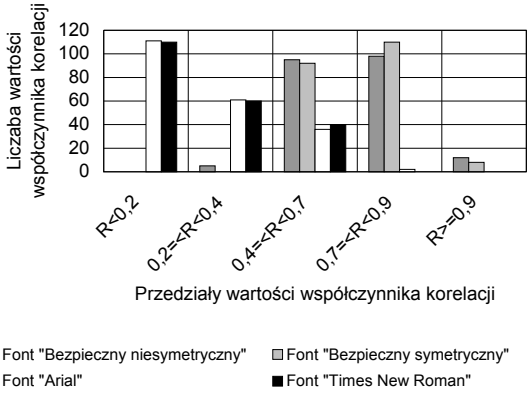
Tabela 2. Liczba znaków, dla których wartości współczynnika korelacji R^Z między znakami małych liter na wyjściu KPI mieści się w zadanych przedziałach

Wartość współczynnika korelacji	Liczba znaków małych liter danego fontu, dla których wartość współczynnika korelacji R^Z (podobieństwa do innego znaku danego fontu) mieści się w zadanym przedziale			
	BN	BS	Arial	Times New Roman
$R < 0.2$	0	0	111	110
$0.2 \le R < 0.4$	5	0	61	60
$0.4 \le R < 0.7$	95	92	36	40
$0.7 \le R < 0.9$	98	110	2	0
$R \ge 0.9$	12	8	0	0

BN – font „Bezpieczny Niesymetryczny", BS – font „Bezpieczny Symetryczny"

Odpowiedzi należy szukać w wynikach analiz sygnałów emisji ujawniających i uzyskiwanych na ich podstawie obrazach. Analogiczne obliczenia R^Z jak dla znaków pierwotnych, przeprowadzono dla znaków zawartych we wspomnianych obrazach.

Otrzymane wyniki potwierdzają słuszność proponowanych kształtów znaków. Wartości współczynnika korelacji R^Z z przedziałów $[0;0{,}2)$ oraz $[0{,}2;0{,}4)$ najczęściej przyjmowane są przez znaki fontów tradycyjnych. Z kolei wartości z przedziałów $[0{,}4;0{,}7)$ oraz $[0{,}7;0{,}9)$ najczęściej przyjmowane są przez znaki fontów bezpiecznych. Oznacza to, że identyfikacja znaków na wyjściu KPI metodą korelacji dwuwymiarowej może być nieskuteczna dla fontów bezpiecznych.



Rys.9. Graficzne przedstawienie liczby znaków fontu (małych liter), dla którego wartość współczynnika korelacji na wyjściu KPI zawiera się w zadanym przedziale wartości R^Z

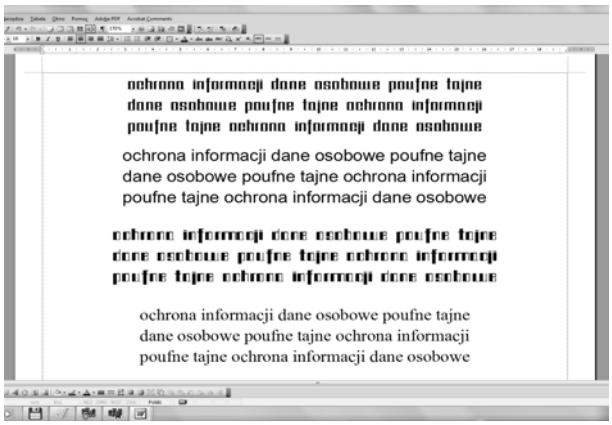
Proces infiltracji elektromagnetycznej

Istotą tworzenia fontów bezpiecznych jest ochrona informacji przed elektromagnetycznym przenikaniem. Sam font w proponowanej postaci nie jest wstanie zapewnić pełnej ochrony. Związane jest to bezpośrednio z powstającymi sygnałami emisji ujawniających o poziomach umożliwiających identyfikację tych emisji, czyli stwierdzeniem, że urządzenie jest źródłem emisji skorelowanych z przetwarzaną informacją. Zatem, fonty bezpieczne nie eliminują całkowicie emisji ujawniających, ale skutecznie uniemożliwiają przeprowadzenie procesu odtworzenia informacji.

Przeprowadzona analiza, przede wszystkim dotycząca zmienności wartości współczynników korelacji znakowej R^Z pokazuje, że założenie o maksymalnym podobieństwie między znakami w ramach danego fontu bezpiecznego zostało spełnione. W większości przypadków wartości te mieszczą się w przedziale $[0{,}7;1]$, co świadczy o wysokim stopniu podobieństwa. Tak wysoka wartość parametru R^Z powinna przełożyć się na trudności w odtworzeniu danych przedstawianych na odtwarzanych obrazach z sygnałów emisji ujawniających.

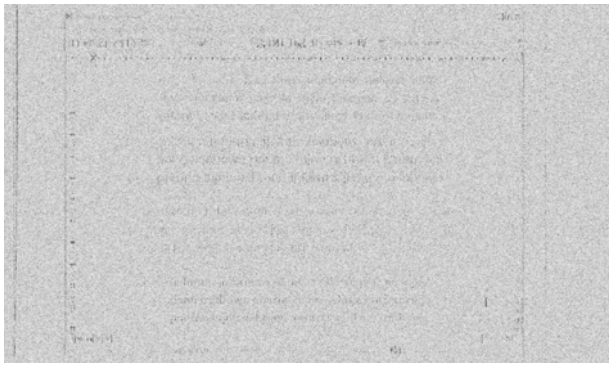
Dla potwierdzenia odporności fontu bezpiecznego na proces infiltracji elektromagnetycznej, przeprowadzono badania w zakresie pomiaru emisji elektromagnetycznych od zestawu komputerowego. Badania wykonano dla częstotliwości od pojedynczych MHz do 1GHz. Jako obraz pierwotny wykorzystano obraz przedstawiony na rysunku 10, zawierający tekst pisany fontem bezpiecznym jak i, dla łatwości prowadzenia analiz porównawczych, fontami tradycyjnymi typu „Arial" oraz „Times New Roman".

Podjęto również próby odtworzenia informacji cyfrowymi metodami przetwarzania obrazów. Wykorzystując te metody, najczęściej zmodyfikowane na potrzeby procesu infiltracji elektromagnetycznej, można na tyle poprawić jakość odtwarzanych obrazów, że możliwy staje się odczyt zapisanego tekstu (w przypadku sygnału emisji ujawniającej o wystarczającym poziomie).

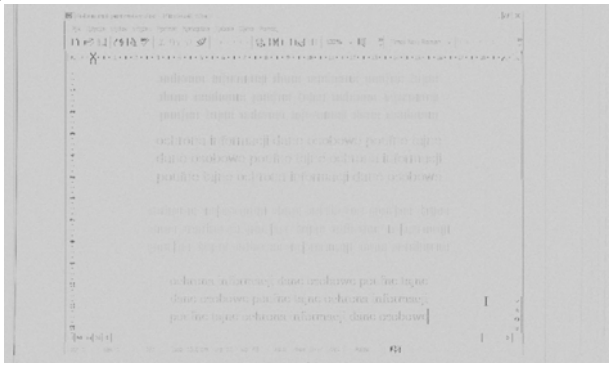


Rys.10. Obrazy pierwotne wykorzystywane w badaniach skuteczności fontu bezpiecznego w ochronie elektromagnetycznej przetwarzanych informacji

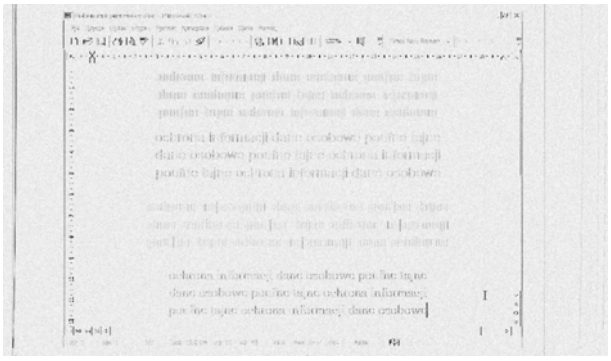
a) obraz uzyskany z sygnału emisji ujawniającej



b) zsumowanie 50 obrazów



c) progowanie wartości amplitud pikseli obrazu uzyskanego po operacji sumowania



Rys.11. Obraz (a) uzyskany z sygnału emisji ujawniającej występującej na częstotliwości 740MHz oraz obrazy przetworzone po zastosowaniu metody sumowania (b) oraz progowania wartości amplitud pikseli (c) – inwersje obrazów

Dotyczy to jednak tylko przypadków wykorzystania fontów tradycyjnych w procesie edycji tekstu. Dla fontów bezpiecznych proces poprawy jakości nie prowadzi do możliwości bezpośredniego odczytania danych. Operacje poprawy jakości obrazów bazujące na metodach cyfrowego ich przetwarzania przeprowadzono dla fontów bezpiecznych jak i dla fontu „Arial" oraz „Times New Roman".

Rejestrowany sygnał emisji ujawniającej odbierany był na częstotliwości 740MHz. Jakość sygnału emisji ujawniającej, z którego metodą rastrowania odtwarzano obraz odpowiadający obrazowi wyświetlanemu na monitorze badanym, uniemożliwiała bezpośredni odczyt zawartych w nim informacji (rys.11).

W pierwszym etapie poprawy jakości obrazu, posłużono się metodą sumowania obrazów, w wyniku której pojawiły się w obrazie poszukiwane elementy. Następnie wykorzystano metodę progowania wartości amplitud pikseli obrazu, która pozwoliła na jeszcze większe wyeksponowanie poszukiwanych danych. Jednak odczyt danych możliwy był jedynie w przypadku gdy ich zapis po stronie pierwotnej odbywał się fontem „Arial" i „Times New Roman". Dane generowane przez fonty „Bezpieczny Symetryczny" jak i „Bezpieczny Niesymetryczny" po stronie odbiorczej nie były czytelne, przez co informacja nimi zapisana stawała się bezpieczna.

Analizując wyżej przedstawione obrazy przede wszystkim można wnioskować, że poziom sygnału emisji ujawniającej jest bardzo niski. W obrazie bez przekształceń słabo zauważalne są elementy tekstu pisane zarówno fontem bezpiecznym jak i fontami tradycyjnymi. Jednak proces infiltracji elektromagnetycznej nie kończy się tylko na etapie odtworzenia obrazu. Obraz taki może być poddawany odpowiedniemu przetwarzaniu zmierzającemu do poprawy jego jakości na tyle, aby możliwym stało się odczytanie zawartych w nim informacji. Wykorzystane dwie metody okazują się skuteczne, ale tylko dla znaków pisanych fontami tradycyjnymi.

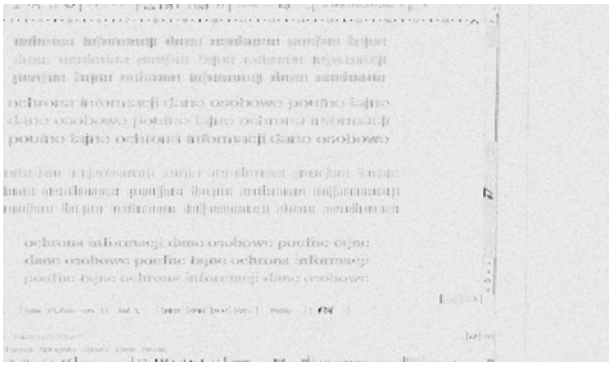
Filtracja logiczna

Wykonywane przez filtry logiczne przekształcenia wartości amplitud pikseli obrazów, można porównać do operacji progowania wartości amplitud. W przypadku filtrów logicznych wymagane jest spełnienie odpowiedniego warunku bazującego na wartościach amplitud pikseli p znajdujących się w określonym otoczeniu rozpatrywanego piksela p_w.

Do analizy obrazów uzyskiwanych z sygnałów emisji ujawniających najczęściej wykorzystywany jest logiczny filtr poziomy. Spowodowane jest to wpływem właściwości KPI na postać sygnału budującego obraz. Działanie filtra oparte jest na analizie wartości amplitud pikseli usytuowanych po lewej i prawej stronie piksela analizowanego [1]. Jeżeli przyjmiemy, że n oznaczać będzie numer kolejnego wiersza obrazu, a m – numer kolejnej kolumny obrazu, wówczas określenie wartości amplitudy piksela $p_w(n,m)$ obrazu przekształconego, odbywa się poprzez analizę i spełnienie odpowiedniego warunku przez wartości amplitud pikseli $p(n,m\text{-}1)$ i $p(n,m\text{+}1)$ obrazu analizowanego (pierwotnego). Zapis analityczny warunku przedstawia się następująco:

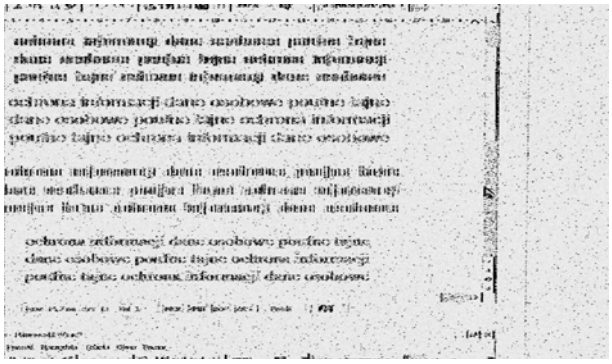
$$p_w(n,m) = \begin{cases} 255 & dla \quad A > B \\ p(n,m) & dla \quad A \le B \end{cases} \tag{7}$$

gdzie: $A = |p(n,m\text{-}1) - p(n,m\text{+}1)|$, B – zadany próg dla danego obrazu wyznaczony metodą „prób i błędów" dla jak najlepszego uzyskania efektu działania filtra.



Rys.12. Obraz odtworzony z sygnału emisji ujawniającej (inwersja obrazu) mierzonej na częstotliwości 740MHz, zawierający tekst pisany (od góry) fontem „Bezpiecznym Symetrycznym", „Arialem", „Bezpiecznym Niesymetrycznym" i „Times New Roman"

Zastosowanie filtracji logicznej do konkretnego przypadku (rys.13) pokazuje, że o ile znaki fontów „Arial" i „Times New Roman" są rozpoznawalne, to znaki fontów bezpiecznych pozostają nadal nieczytelne.



Rys.13. Obraz (inwersja obrazu) z rys.12 poddany procesowi dwukrotnej filtracji logicznym filtrem poziomym o wartości progu B równym 35

Operacje na histogramach

Obrazy uzyskiwane z sygnałów emisji ujawniających bardzo często, oprócz wielu zakłóceń, są zbyt ciemne mimo wykorzystania pełnej dynamiki układów próbkujących. Celem wydobycia z nich poszukiwanych danych (cech dystynktywnych ułatwiających podejmowanie słusznych decyzji) muszą one zostać poddane odpowiednim przekształceniom m.in. poprzez rozszerzanie lub wyrównywanie histogramów wartości amplitud pikseli, pozwalającym na ich analizę w zakresie rozpoznania znaków graficznych.

Metody jakimi się posłużono to rozszerzenie histogramu z wykorzystaniem odcinków liniowych, rozszerzenie wybranego zakresu histogramu oraz wyrównanie histogramu bazujące na tablicy transformacji, w literaturze powszechnie oznaczane przez *LUT* [1] (ang. Look Up Table).

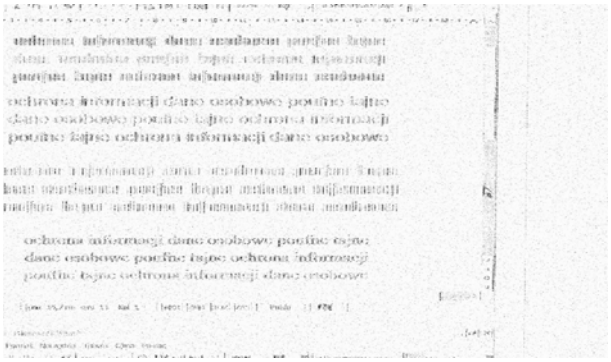
rozszerzenie histogramu (odcinków liniowych)

Rozszerzanie histogramu z wykorzystaniem odcinków liniowych pozwala na dowolną zmianę jasności pikseli obrazu w zależności od nachylenia odcinków względem osi odciętej. W tym przypadku nowe wartości pikseli $p_w(n,m)$ mogą zostać skupione lub rozszerzone na skali odcieni szarości P_w w stosunku do wartości amplitud pikseli pierwotnych $p(n,m)$ znajdujących się na skali odcieni szarości P. W zależności od przyjętych wartości p_1, p_2, p_{w1} i p_{w2} otrzymujemy różnorodność nachyleń odcinków transformacji. Nowe wartości jasności P_w pikseli $p_w(n,m)$ opisywane są zależnością:

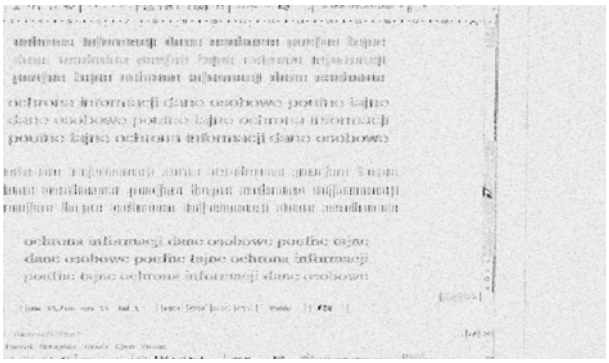
$$p_w(n,m)=\begin{cases} a_1 \cdot p(n,m)+b_1 & dla \quad C \\ a_2 \cdot p(n,m)+b_2 & dla \quad D \\ a_3 \cdot p(n,m)+b_3 & dla \quad E \end{cases} \tag{8}$$

gdzie: $C \rightarrow p_{\min} \le p(n,m) < p$, $D \rightarrow p_1 \le p(n,m) < p_2$, $E \rightarrow p_2 \le p(n,m) \le p_{\max}$, $a_1 = p_{w1}/p_1$, $b_1 = 0$, $a_2 = (p_{w2}-p_{w1})/(p_2-p_1)$, $b_2 = p_{w1}-a_2 \cdot p_1$, $a_3 = (p_{wmax}-p_{w2})/(p_{max}-p_2)$, $b_3 = p_{w2}-a_3 \cdot p_2$, a_1, a_2, a_3 – są współczynnikami nachylenia.

Dobierając odpowiednie wartości p_1, p_2, p_{w1} i p_{w2}, najczęściej metodą prób i błędów, uzyskujemy różne postacie obrazów zawierających mniej lub bardziej czytelne poszukiwane dane (rys.14 i rys.15).



Rys.14. Obraz (inwersja obrazu) z rysunku 12 poddany procesowi rozszerzenie histogramu z wykorzystaniem odcinków liniowych o wartościach parametrów: p_1 =30, p_{w1} = 5, p_2 = 110, p_{w2} = 250

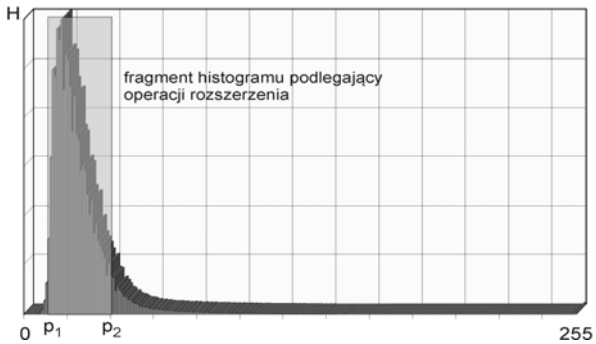


Rys.15. Obraz (inwersja obrazu) z rysunku 12 poddany procesowi rozszerzenie histogramu z wykorzystaniem odcinków liniowych o wartościach parametrów: p_1 = 25, p_{w1} = 20, p_2 = 125, p_{w2} = 230

Zastosowanie operacji rozszerzenia histogramu wartości amplitud pikseli obrazu z rysunku 12, nie spowodowało poprawy czytelności znaków zapisanych fontami bezpiecznymi. Znaki fontów „Arial" oraz „Times New Roman" stały się bardziej kontrastowe. Znaki fontów bezpiecznych nadal przedstawione są w postaci „paczek" linii pionowych bardziej lub mniej kontrastowych, uniemożliwiających jednak ciągle odczyt znaków i informacji nimi zapisanych.

rozszerzenie wybranego zakresu histogramu

W przypadku obrazów emisji ujawniającej bardzo często istotny jest fragment histogramu. Wynika to z faktu, że poszukiwane dane ukryte w obrazie niewiele różnią się od otaczającego tła. Z tym związana jest bliskość wartości amplitud pikseli niosących poszukiwaną informację z wartościami amplitud pikseli tła. Z tego względu operacji rozszerzania histogramu powinien być poddawany jedynie jego fragment (rys.16).



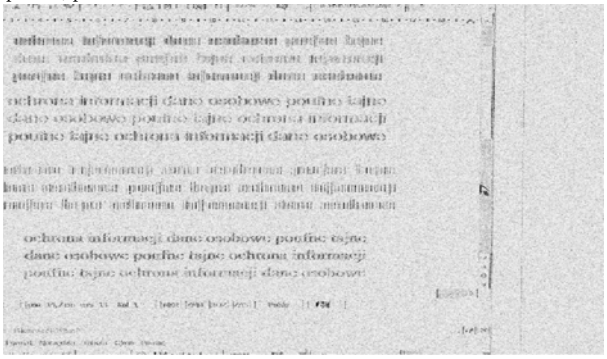
Rys.16. Rozszerzanie wybranego zakresu histogramu obrazu odtworzonego

Jest to metoda mniej złożona niż rozszerzanie histogramu z wykorzystaniem odcinków liniowych ale równie skuteczna w procesie infiltracji elektromagnetycznej. Wymaga ona podania jedynie wartości dwóch parametrów p_1 i p_2 (rys.16). Wówczas:

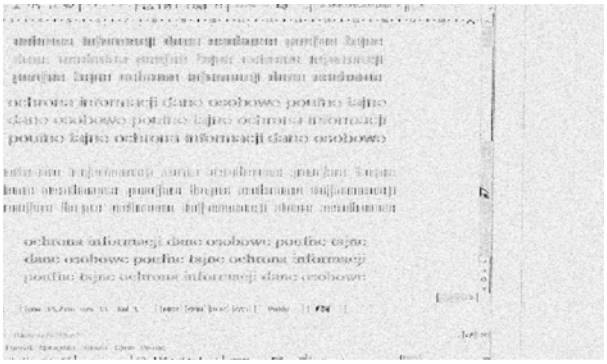
$$p_w(n,m)=\begin{cases}\dfrac{p(n,m)-p_1}{p_2-p_1}\cdot 255 & dla \quad D\\ 0 & dla \quad F \text{ lub } G\end{cases}, \tag{9}$$

gdzie: $F \rightarrow p(n,m) < p_1$, $G \rightarrow p(n,m) > p_2$, p_1 – minimalna wartość amplitudy piksela obrazu odtworzonego dla rozpatrywanego fragmentu histogramu, p_2 – maksymalna wartość amplitudy piksela obrazu dla rozpatrywanego fragmentu histogramu, $p(n,m)$ – wartość amplitudy piksela o współrzędnych n i m obrazu odtworzonego, $p_w(n,m)$ – wartość amplitudy piksela o współrzędnych n i m obrazu przekształconego.

a) $p_1 = 9$, $p_2 = 115$



b) $p_1 = 12$, $p_2 = 140$



Rys.17. Obraz (inwersja obrazu) uzyskany w wyniku rozszerzenia fragmentu histogramu obrazu z rysunku 12

Stosując operację rozszerzenia histogramu opisaną zależnością (9) dla obrazu z rysunku 12, przy zadanych wartościach progowych p_1 i p_2, uzyskano obraz przedstawiony na rysunku 17.

wyrównanie histogramu (tablica transformacji LUT)

Tablica transformacji *LUT* wykorzystywana w procesie wyrównywania histogramów wartości amplitud pikseli obrazów opisywana jest zależnością:

$$LUT_u = \frac{D_u - D_{\min}}{1 - D_{\min}} \cdot (L-1), \tag{10}$$

gdzie: $D_{\min}$ jest pierwszą niezerową wartością dystrybuanty określonej zależnością:

$$D_u = \sum_{j=0}^{u} H_{j\,unorm}\ , \tag{11}$$

oraz $H_{j\ unorm}$ – unormowany histogram wartości amplitud pikseli analizowanego obrazu, $u = 0, 1, ..., q, ..., L–1$, $L–1 = 255$, $D_{\max} = D_{L-1}$.

Praktyka jednak pokazuje [1], że transformacja z wykorzystaniem typowych tablic *LUT* nie zawsze jest zadawalająca. Powodem jest niewystarczający kontrast obrazu. Wartości amplitud pikseli znaków graficznych porównywalne są z wartościami amplitud tła zakłóceń. Rozwiązaniem jest usunięcie z obrazów tych wartości amplitud pikseli, które decydują o jasności tła. W tym celu stosuje się modyfikację tablic *LUT* opartą na analizie zmienności wartości dystrybuanty D rozkładu prawdopodobieństwa wartości amplitud pikseli obrazu analizowanego. Wartość dystrybuanty D równa 0.88 [1] określa warunek, przy spełnieniu którego po dokonaniu wyrównania histogramu obrazu odtworzonego (analizowanego), jakość (kontrast) uzyskiwanego obrazu pozwala na lepsze rozróżnianie znaków graficznych niż w przypadku typowego rozszerzenia histogramu. Wówczas:

$$D_{\min} = D_u \tag{12}$$

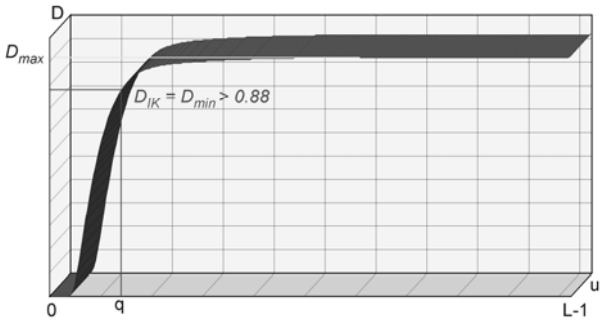
gdzie: $u = 0, 1, ..., q$; q – pierwsza wartość amplitudy piksela obrazu odtworzonego, dla której D_{IK}, określone zależnością (13), osiągnie wartość większą od $0{,}88$ (rys.18). Dalsze zwiększanie wartości q może zwiększyć D_{IK}, ale może pogorszyć czytelność obrazu, oraz:

$$D_{IK} = D_{\min}/D_{\max} > 0{,}88. \tag{13}$$

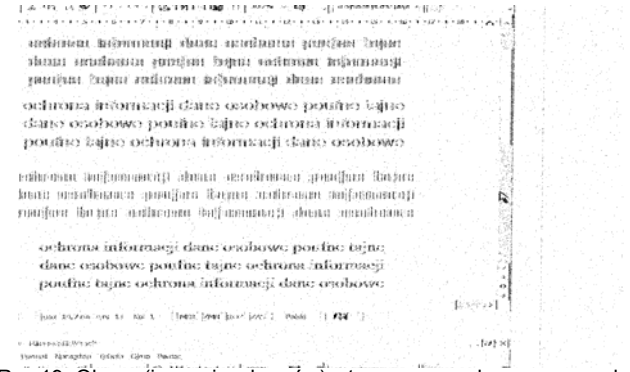
Ponieważ $D_{\max} = 1$, to:

$$D_{IK} = D_{\min} > 0{,}88. \tag{14}$$

Określona w ten sposób wartość q, pozwala na obliczenie wartości współczynników tablicy *LUT* dla zawężonego zakresu skali odcieni barw szarości, czyli dla $u = q, ..., L–1$.



Rys.18. Przykładowy przebieg dystrybuanty rozkładu prawdopodobieństwa z zaznaczonymi wartościami $D_{\min}$ i $D_{\max}$ określającymi wartość kryterium D_{IK}

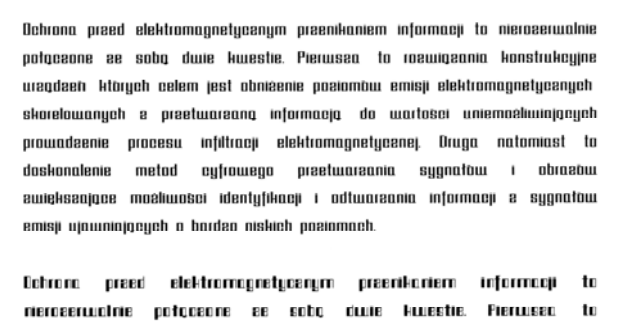


Rys.19. Obraz (inwersja obrazów) otrzymany z obrazu z rysunku 12 po przeprowadzeniu operacji wyrównania histogramu bazującej na zmodyfikowanych wartościach tablicy *LUT* ($q = 51$)

Analizując powyższe próby modyfikacji obrazów zmierzające do umożliwienia odczytu danych zapisanych fontami bezpiecznymi zauważamy, że jakość obrazów poprawia się, przez co stopień percepcji tekstu zapisanego fontami „Arial" i „Times New Roman" wzrasta, natomiast dane zapisane fontami bezpiecznymi pozostają nieczytelne. Potwierdzona tym samym zostaje słuszność stosowania fontów specjalnie zaprojektowanych, które mają wspomagać ochronę danych przed przenikaniem elektromagnetycznym.

Próby wykorzystania programów OCR

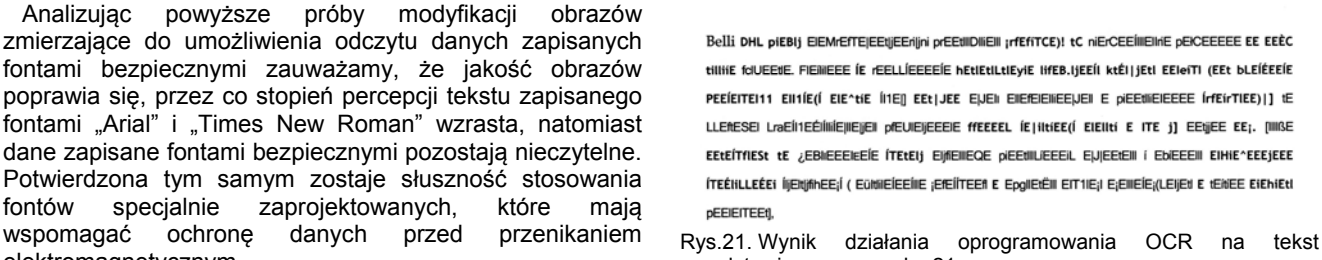
Może nie najważniejszym ale interesującym zagadnieniem ochrony informacji jest bezpieczeństwo związane z brakiem możliwości wykorzystania programów OCR do dokumentów w formie papierowej. Dotyczy to sytuacji kiedy podejmowane są próby kopiowania dokumentów do postaci elektronicznej umożliwiającej jego przetwarzanie, którego jednym z elementów jest dystrybucja dokumentu. Nie dotyczy to przypadków kopiowania dokumentów do postaci tzw. skanów, które uniemożliwiają modyfikowanie treści dokumentu.



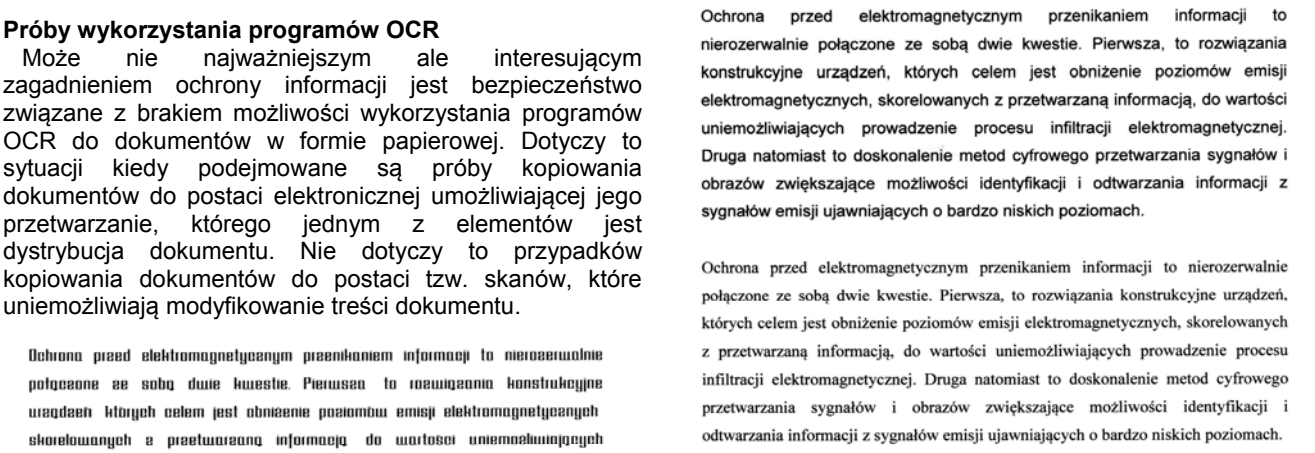
Rys.20. Tekst podlegający procesowi optycznego rozpoznawania znaków pisany fontem „Bezpieczny Symetryczny" i „Bezpieczny Niesymetryczny"

Zaprojektowane fonty bezpieczne wykazują, obok odporności na proces infiltracji elektromagnetycznej, odporność na działanie programów OCR. Skanowanie treści dokumentu pisanego tymi fontami skutkuje otrzymaniem dokumentu o treści niezrozumiałej dla człowieka. Otrzymywany ciąg znaków nie buduje czytelnych wyrazów a tym samym zdań, które zawierałyby informację skorelowaną z niejawnym dokumentem pierwotnym.

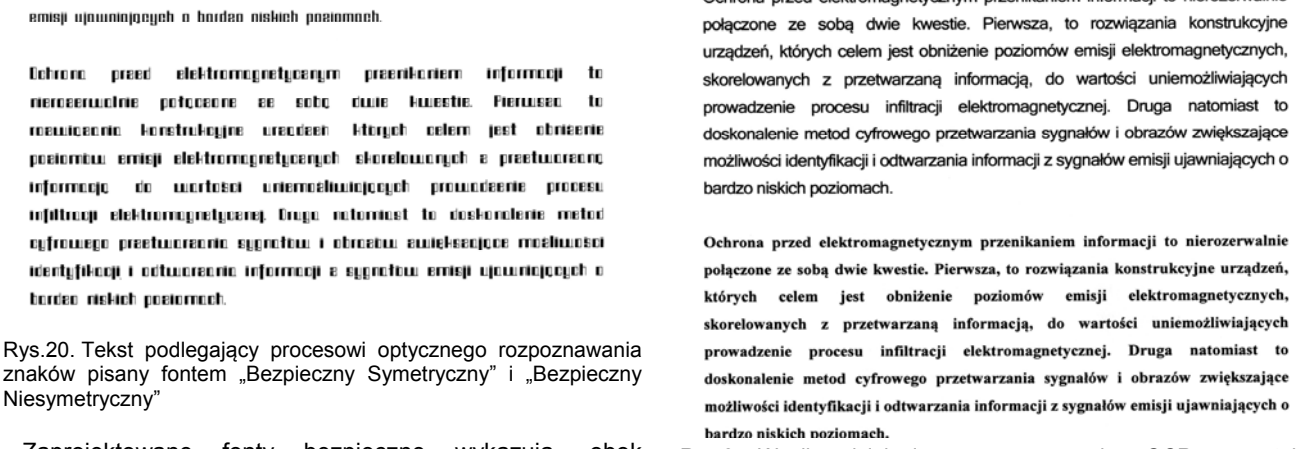
Procesowi optycznego rozpoznawania znaków poddano dokument przedstawiony na rysunku 20. Zawiera on tekst pisany zarówno fontem „Bezpiecznym Symetrycznym" jak i „Bezpiecznym Niesymetrycznym". Wyniki dotyczą przypadku gdy nie prowadzono zabiegów „uczenia" oprogramowania rozpoznawania znaków.



Rys.21. Wynik działania oprogramowania OCR na tekst przedstawiony na rysunku 21



Rys.22. Tekst podlegający procesowi optycznego rozpoznawania znaków pisany fontem „Arial" oraz „Times New Roman"



Rys.24. Wynik działania oprogramowania OCR na tekst przedstawiony na rysunku 23

W przypadku typowych fontów „Arial" i „Times New Roman" stosowanych na co dzień przez nas wszystkich, optyczne rozpoznawanie znaków zostało przeprowadzone z

pełnym sukcesem. Brak jest fałszywych decyzji, a tekst jest wiernie przeniesiony do postaci edytowalnej.

Zgoła odmienna sytuacja ma miejsce w przypadku próby optycznego rozpoznawania znaków pisanych fontem „Bezpieczny Symetryczny" i „Bezpieczny Niesymetryczny". Podejmowane decyzje związane z rozpoznaniem pojedynczych znaków nie mają żadnego związku ze znakiem pierwotnym. Otrzymany tekst jest zbiorem losowych znaków, które nie tworzą logicznego ciągu. Informacja zapisana w tej postaci jest niezrozumiała i nie stanowi podstaw do dalszego przetwarzania mającego na celu uzyskanie dokumentu zbieżnego z dokumentem skanowanym.

Podsumowanie

W artykule przedstawiono trzy rodzaje fontów bezpiecznych: „Bezpieczny Symetryczny", „Bezpieczny Niesymetryczny" i „Bezpieczny Prosty". Nazwy fontów bezpośrednio wiążą się z wyglądem znaków danego fontu. Font „Bezpieczny Symetryczny" to pogrubione elementy, prawy i lewy, każdego znaku tak aby upodobnić je jak najbardziej do siebie. Font „Bezpieczny Niesymetryczny" to z kolei pogrubiona lewa część każdego znaku. Font „Bezpieczny Prosty" to typ kroju jednoelementowego o jednakowej szerokości linii pionowych i poziomych. Proces upodobnienia znaków w ramach każdego fontu spowodował, że współczynniki korelacji między poszczególnymi znakami w większości przypadków mieszczą się w przedziale od $0{,}7$ do $1{,}0$. Jest to wartość bardzo duża mówiąca o dość silnej i bardzo silnej zależności między znakami. W przypadku fontów „Arial" oraz „Times New Roman" wartości te są znacznie mniejsze.

Wysoki stopień podobieństwa między znakami powoduje, że odtworzenie informacji z rejestrowanych sygnałów emisji ujawniającej jest bardzo utrudniony i skutkuje to zjawiskiem podejmowania wielu fałszywych decyzji w procesie rozpoznawania znaków, w którym wykorzystywana jest metoda korelacji dwuwymiarowej. Odtwarzanie informacji metodą wzrokową jest praktycznie niemożliwe.

Font bezpieczny chroni informacje niejawne nie tylko przed elektromagnetycznym przenikaniem. Jednym z celów stworzenia fontu bezpiecznego było również uodpornienie go na działanie programów OCR. Cel został osiągnięty. Zastosowanie oprogramowania OCR do tekstu napisanego i wydrukowanego wspomnianym fontem powoduje, że postać rozpoznanego tekstu całkowicie nie odpowiada tekstowi pierwotnemu, przez co nie nadaje się do dalszej edycji. W uzyskanym obrazie występują znaki, które nie mają żadnego związku ze znakami występującymi w tekście skanowanym. To zjawisko powoduje, że nie jest możliwe przeniesienie wprost wydrukowanego fontem bezpiecznym tekstu do formy edytowalnej i dystrybuowanie go w formie elektronicznej.

LITERATURA

[1] Kubiak I., *Metody analizy i cyfrowego przetwarzania obrazów w procesie infiltracji elektromagnetycznej*, Wydawnictwo Wojskowej Akademii Technicznej 2013, ISBN 978-83-62954-86-5 (monografia);

[2] Kubiak I., Przybysz A., Musiał S., Grzesiak K., *Elektromagnetyczne bezpieczeństwo informacji*, Wydawnictwo Wojskowej Akademii Technicznej 2009, ISBN 978-83-61486-32-9 (monografia);

[3] Kubiak I., Przybysz A., Musiał S., Grzesiak K., *Generator rastra w procesie infiltracji elektromagnetycznej*, Wydawnictwo Wojskowej Akademii Technicznej 2012, ISBN 978-83-62954-28-5 (monografia);

[4] Tomasz P. Zieliński, *Cyfrowe przetwarzanie sygnałów. Od teorii do zastosowań*, WKŁ, 2009;

[5] 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;

[6] 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;

[7] 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;

[8] Grzesiak K., Przybysz A., *Emission security of laser printers*, MCC 2010: Military Communications and Information Systems Conference, Wrocław 2010;

Autor: *dr inż. Ireneusz Kubiak, Wojskowy Instytut Łączności, ul. Warszawska 22a, 05-130 Zegrze Południowe, E-mail: i.kubiak@wil.waw.pl.*