



Metody usuwania szumu z obrazów twarzy

dr inż. TOMASZ ANDRYSIAK ¹⁾, mgr inż. LECH NAUMOWSKI ²⁾

¹⁾ Instytut Telekomunikacji, Uniwersytet Technologiczno-Przyrodniczy w Bydgoszczy

²⁾ Instytut Maszyn Matematycznych, Warszawa

Obecnie biometryczne metody identyfikacji i rozpoznawania osób na podstawie obrazów twarzy zyskują coraz większą popularność, dzięki dynamicznie rozwijającym się systemom komputerowym oraz techniką obrazowania. Rosnące praktyczne ich znaczenie spowodowane jest głównie poprzez łatwość pozyskiwania oraz brak interakcji systemu z osobą której próbki biometryczne są pobierane. Dlatego też systemy te są szczególnie czułe na warunki akwizycji oraz jakość eksploatowanych urządzeń obrazowania. Istotnymi czynnikami mającymi wpływ na przedmiotową jakość są warunki oświetlenia, zniekształcenia geometryczne systemów akwizycji oraz szumy wprowadzane przez detektory optyczne i kanały transmisyjne.

Szum jest podstawowym czynnikiem utrudniającym dalszą analizę obrazu. Potrafi w sposób znaczący wpłynąć na jego interpretację i rozpoznanie. Dlatego też ważnym zagadnieniem jest rozwijanie metod pozwalających redukować jego wpływ na przetwarzane obrazy w systemach biometrycznych. Metody realizujące to zadanie najczęściej wykorzystują różnego typu techniki filtracji, analizy w dziedzinie określonej transformaty oraz dekompozycji obrazów.

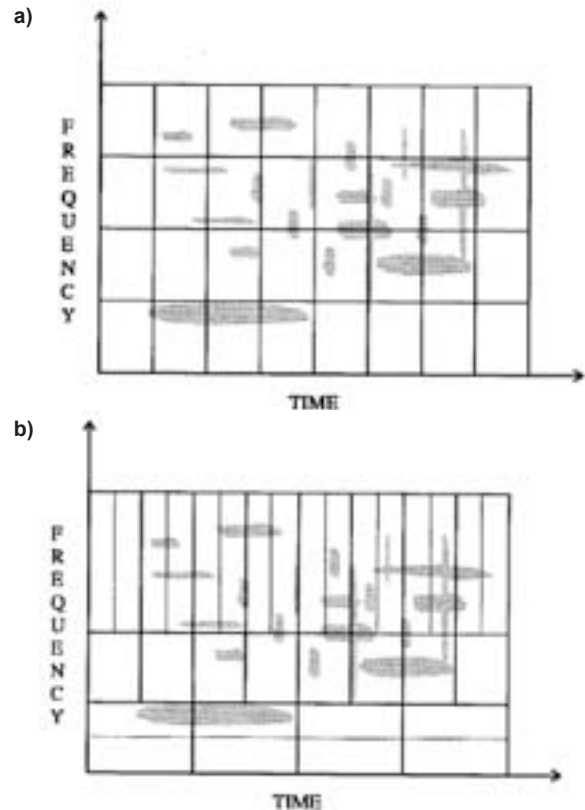
Reprezentacje sygnału s realizowane jako liniowe rozwinięcia:

$$s = \sum_n c_n g_{\gamma_n} \quad (1)$$

względem określonego zbioru funkcji g_{γ} dobrze zlokalizowanych względem czasu i/lub częstotliwości są w wielu przypadkach wystarczająco precyzyjne i optymalne [4]. W zależności od dokonanego wyboru funkcji g_{γ} współczynniki rozwinięcia c_n reprezentują różne cechy sygnału s .

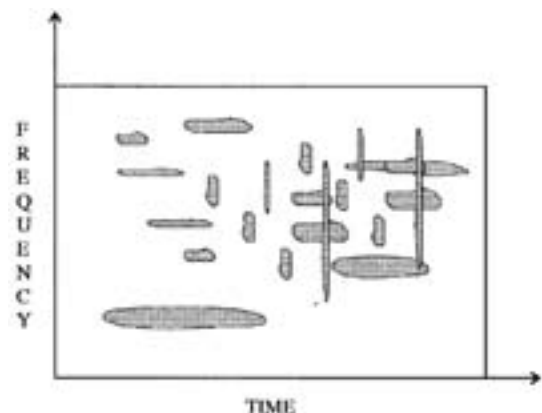
Jeżeli wielkości elementów strukturalnych analizowanego sygnału różnią się w sposób istotny od stałej skalującej funkcji g_{γ} to współczynniki c_n nie stanowią optymalnej reprezentacji przedmiotowego sygnału (rys. 1a). Dlatego też sygnały z elementami o zmiennych długościach wymagają stosowania funkcji bazowych g_{γ} o różnych skalach (rys. 1b). Rozwiązanie takie posiada także istotne ograniczenie związane z powiązaniem parametru częstotliwości i parametru skali dla określonych funkcji bazowych g_{γ} .

W przypadku skomplikowanych struktur sygnałów nie jesteśmy w stanie określić optymalnych parametrów skali i częstotliwości dla określonych funkcji bazowych g_{γ} . Wówczas naturalnym rozwiązaniem jest wprowadzenie bardziej różnorodnych i liczniejszych a także dostosowanych do charakteru sygnału zbiorów funkcji zwanych słownikami z redundancją [4]. Otrzymujemy wówczas reprezentacje znacznie bardziej uniwersalne i elastyczne (rys. 2).



Rys. 1. Symboliczny podział płaszczyzny czasowo-częstotliwościowej dla: a) rozwinięć ortogonalnych DCT, b) diadycznych rozwinięć falkowych Haara [2]

Fig. 1. Symbolic division of the time-frequency plane for: a) orthogonal DCT decomposition, b) orthogonal dyadic Haar decomposition [2]



Rys. 2. Funkcje bezpośrednio dostosowane do sygnału [2]
Fig. 2. Basis functions adapted directly to the signal [2]



Jednakże dekompozycja sygnału względem słownika wymaga wówczas ciągłego wyszukiwania i dopasowywania odpowiednich jego elementów które najlepiej odzwierciedlają pożądane cechy analizowanego sygnału. Dopasowanie to powinno być realizowane w taki sposób, aby maksymalizować korelację pomiędzy wybranym elementem słownika, a pozostałą częścią przetwarzanego sygnału. Istnieje więc potrzeba określenia miary jakości takiego dopasowania oraz algorytmu zapewniającego znalezienie najlepszej w sensie tej miary reprezentacji sygnału [4].

Przedstawioną ideę adaptacyjnej dekompozycji stosowano w analizie czasowo-częstotliwościowej sygnałów akustycznych i biomedycznych [7, 8, 14]. W dziedzinie przetwarzania obrazów cyfrowych stosowano najczęściej w zadaniach związanych z wydzieleniem cech, kompresją oraz metodach poprawy jakości analizowanych obrazów [1–3, 9].

Adaptacyjna dekompozycja obrazu

Dekompozycja obrazu f prowadząca do adaptacyjnej jego reprezentacji poprzez wybór zbioru atomów g_{γ} słownika D sprowadza się do zagadnienia jak najlepszego ich dopasowania względem analizowanego obrazu f tj. minimalizacji błędu aproksymacji δ .

$$\delta = \left\| f - \sum_{i=1}^{M-1} c_i g_{\gamma_i} \right\| \quad (2)$$

Reprezentację optymalną możemy określić jako taki podzbiór elementów słownika, którego liniowa kombinacja tłumaczy największy procent energii sygnału wśród wszystkich podzbiorów o tej samej liczebności. Wybór takiej reprezentacji jest obliczeniowo NP-trudny, toteż w praktyce zadowalamy się iteracyjnym rozwiązaniem adaptacyjnym zwanym algorytmem MP [15]. Wynikiem działania tego algorytmu jest projekcja elementów strukturalnych sygnału na wybrane funkcjesłownika zwane atomami.

Algorytm MP

W celu realizacji dekompozycji obrazu f wyznaczać będziemy liniowe rozwinięcie względem zbioru atomów g_{γ} wybranych ze słownika D tak aby były najlepiej dopasowane do elementów strukturalnych analizowanego obrazu f .

W każdym kroku n algorytmu MP dokonywać będziemy kolejnych dekompozycji obrazu f za pomocą rzutów ortogonalnych na elementy słownika D . Wówczas w n -tym kroku dekompozycji otrzymujemy:

$$R^n f = R^{n-1} f - \left\langle R^{n-1} f, g_{\gamma_{n-1}} \right\rangle g_{\gamma_{n-1}}, \quad (3)$$

gdzie $\langle *, * \rangle$ oznacza iloczyn skalarny a $R^n f$ stanowi residuum będące wynikiem dekompozycji f w kierunku g_{γ} . Dla residuum zerowego rzędu zachodzi oczywista zależność $R^n f = f$.

Wyboru atomu g_{γ} dokonujemy na podstawie:

$$g_{\gamma_n} = \arg \max_{g_{\gamma_i} \in G} \left| \left\langle R^{n-1} f, g_{\gamma_i} \right\rangle \right|, \quad (4)$$

gdzie G jest zbiorem indeksów słownika D .

Dla wybranego za pomocą zależności (4) atomu g_{γ} minimalizujemy residuum w kolejnym kroku n algorytmu MP. Z uwagi na ortogonalności wektorów g_{γ} i $R^n f$ zachodzi związek:

$$\left\| R^{n-1} f \right\|^2 = \left| \left\langle R^{n-1} f, g_{\gamma_{n-1}} \right\rangle \right|^2 + \left\| R^n f \right\|^2. \quad (5)$$

Kontynuując proces dekompozycji do N -tego poziomu otrzymujemy:

$$R^N f = f - \sum_{n=0}^{N-1} \left\langle R^n f, g_{\gamma_n} \right\rangle g_{\gamma_n}. \quad (6)$$

Podobnie do zależności (5) możemy zapisać:

$$\left\| f \right\|^2 = \sum_{n=0}^{N-1} \left| \left\langle R^n f, g_{\gamma_n} \right\rangle \right|^2 + \left\| R^N f \right\|^2, \quad (7)$$

otrzymując równanie zachowania energii [6].

Liczba iteracji N w których dokonujemy dekompozycji residuów zależy od wymaganej dokładności rekonstrukcji obrazu f i dana jest zależnością (8) będącą jednocześnie kryterium stopu algorytmu MP

$$\left\| R^N f \right\|^2 \leq \beta \left\| f \right\|^2, \quad (8)$$

gdzie parametr $\beta \in (0, 1)$.

Szybkość malenia normy residuum zależy od korelacji pomiędzy kolejnymi residuami obrazu, a wybranymi atomami słownika. Jeżeli obraz jest sumą składowych o dużej energii będących atomami słownika, to współczynniki korelacji obrazu i jego residuów są znaczne. Wówczas ich norma szybko maleje ponieważ składowe o dużej energii są elementami strukturalnymi obrazu dobrze skorelowanymi z wybranymi atomami słownika [4, 14].

Słowniki funkcji bazowych

Dokładna reprezentacja analizowanego obrazu w słowniku większym niż baza wprowadza redundancję. Pożądaną zwężoność możemy osiągnąć godząc się na niedokładności rekonstrukcji obrazu, ale z wykorzystaniem możliwie niewielkiej ilości funkcji [7].

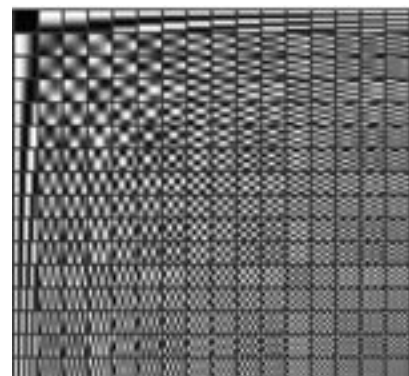
Słownik DCT

Zbiór atomów słownika DCT może otrzymać wykorzystując funkcje bazowe $d_i(x)$ transformacji DCT:

$$d_i(x) = \left(\frac{2}{L} \right)^{\frac{1}{2}} A_i \cdot \cos \left(\frac{\pi x}{2L} (2i+1) \right), \quad (9)$$

gdzie:

$$A_i = \begin{cases} \frac{1}{\sqrt{2}} & \text{dla } i = 0 \\ 0 & \text{dla innych } i \end{cases} \quad (10)$$



Rys. 3. Słownik DCT. Fig. 3. DCT dictionary



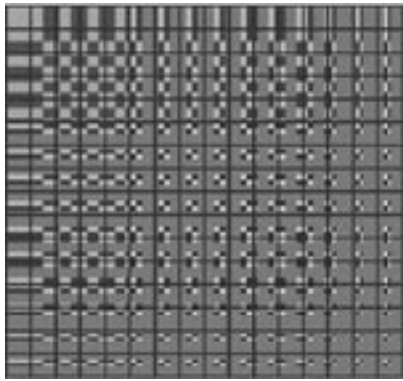
Słownik Haara

Zbiór atomów słownika Haara generujemy poprzez operacje przesunięcia i skalowania funkcji $h(x)$ [12]:

$$h_{mn}(x) = 2^{-\frac{m}{2}} \cdot h(2^{-m}x - n), \quad (11)$$

dla $m, n = \dots, -2, -1, 0, 1, 2, \dots$, gdzie funkcję podstawową Haara $h(x)$ określamy jako:

$$h(x) = \begin{cases} 1 & \text{dla } 0 \leq x < \frac{1}{2} \\ -1 & \text{dla } \frac{1}{2} \leq x < 1 \\ 0 & \text{dla } \text{dla innych } x \end{cases} \quad (12)$$



Rys. 4. Słownik Haara. Fig. 4. Haar dictionary

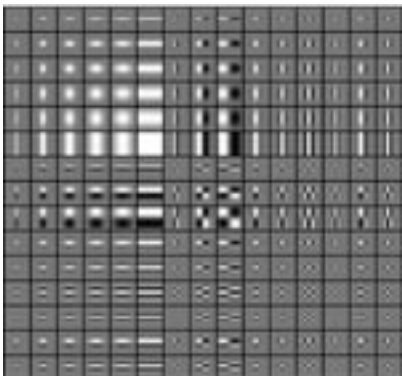
Słownik Gabora

Zbiór atomów słownika Gabora możemy wygenerować poprzez skalowanie, przesunięcie i modulację gaussowskiej funkcji okna $g(x)$ [11,13]. Definiujemy wówczas zbiór $Z = R^+ \times R^2$, $\gamma_n \in Z$ indeksów atomów $\gamma_n = \{u_n, \vartheta_n, \varphi_n\}$, gdzie parametr u odpowiedzialny jest za skalowanie, ϑ jest przesunięciem, a φ częstotliwością modulującą [15].

Otrzymujemy wówczas:

$$g_{\gamma_n}(x) = \frac{K_{\gamma_n}}{\sqrt{u_n}} g\left(\frac{x\vartheta_n}{u_n}\right) \cos(x\vartheta + \varphi_n), \quad (13)$$

gdzie $\vartheta_n \neq 0$, faza $\varphi_n \in (0, 2\pi)$, a stała K_{γ_n} jest tak dobrana aby $\|g_{\gamma_n}\| = 1$.



Rys. 5. Słownik Gabora. Fig. 5. Gabor dictionary

Parametry dla atomów Gabora [2]
Parameters for the Gabor atoms [2]

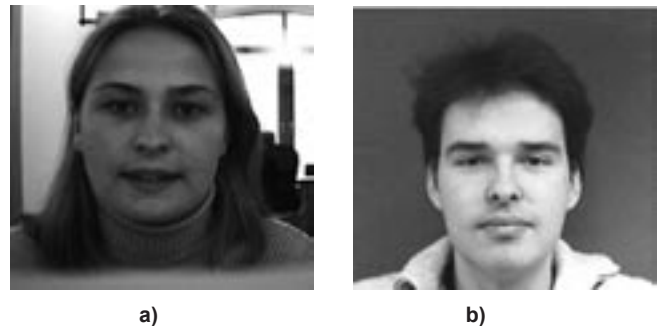
n	u	ϑ	φ
1	2	0	0
2	3	0	0
3	4	0	0
4	5	0	0
5	6	0	0
6	8	0	0
7	10	0	0
8	11	0	0
9	1	1	π/2
10	5	1	π/2
11	11	1	π/2
12	10	3	0
13	8	2	0
14	4	2	0
15	4	2	π/4
16	6	4	π/4

W tabeli przedstawiono wartości parametrów (u, ϑ, φ) wykorzystywanych w procesie tworzenia atomów słownika Gabora na podstawie zależności (13).

Na rysunku 5 przedstawiono obraz słownika Gabora modelowanego za pomocą parametrów określonych w tabeli.

Wyniki eksperymentów

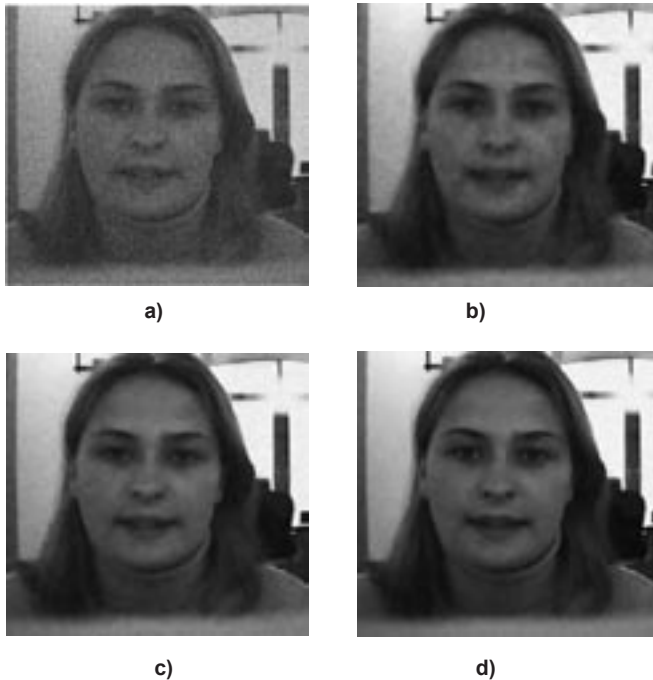
W przeprowadzonych badaniach analizowano obrazy testowe twarzy zaszumiane szumem gaussowskim o zerowej wartości średniej i zmiennym odchyleniu standardowym $\sigma \in (0, 100)$. Proces usuwania szumu polegał na selekcji współczynników dekompozycji realizowanej poprzez kryterium stopu algorytmu MP zgodnie z zależnością (8) dla parametru $\beta = 0.2$. Wszystkie wykorzystywane słowniki składały się z 16×16 atomów, a każdy pojedynczy atom był blokiem o rozmiarze 8×8 elementów.



Rys. 6. Obrazy testowe: a) BioID_0437, b) IMM 07-1
Fig. 6. Sample images: a) BioID_0437, b) IMM 07-1

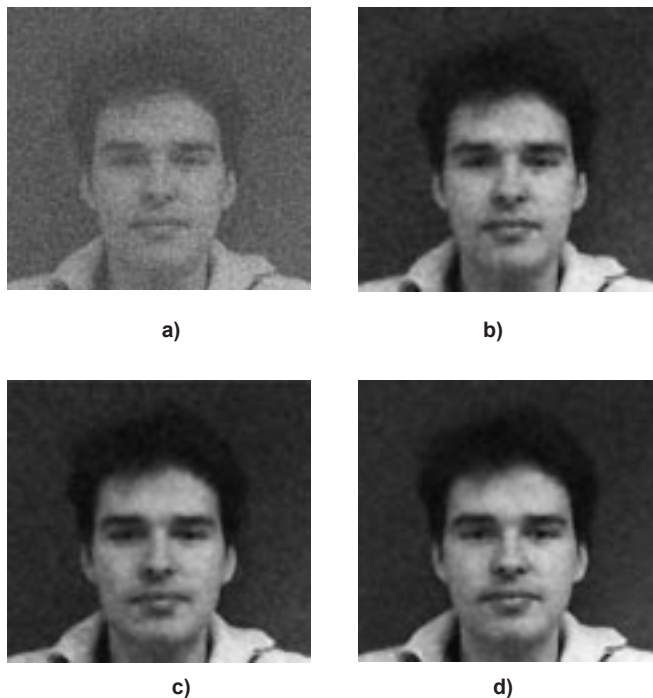
Na rysunku 6 przedstawiono oryginalne obrazy testowe wybrane z baz IMM oraz BioID wykorzystywane w celu zilustrowania przykładowych wyników prezentowanej metody usuwania szumów z przetwarzanych obrazów testowych [18, 20].

Przykładowe obrazy będące efektem opisywanej metody usuwania szumów przedstawiono na rys. 7 i 8. Obrazują one dobrą subiektywną ocenę prezentowanej metody szczególnie dla parametru $\sigma = 30$ (rys. 7) choć analizując szczegóły



Rys. 7. Przykładowe rezultaty usuwania szumu dla obrazu BioID_0437: a) obraz zaszumiany $\sigma = 30$, PSNR = 18,58 dB, b) Haar PSNR = 32,51 dB, c) DCT PSNR = 33,63 dB, d) GaborPSNR = 34,16 dB.

Fig. 7. Example of the denoising results for the image BioID_0437: a) noisy image $\sigma = 50$, PSNR = 18,58 dB, b) Haar PSNR = 32,51 dB, c) DCTPSNR = 33,63 dB, d) GaborPSNR = 34,16 dB.

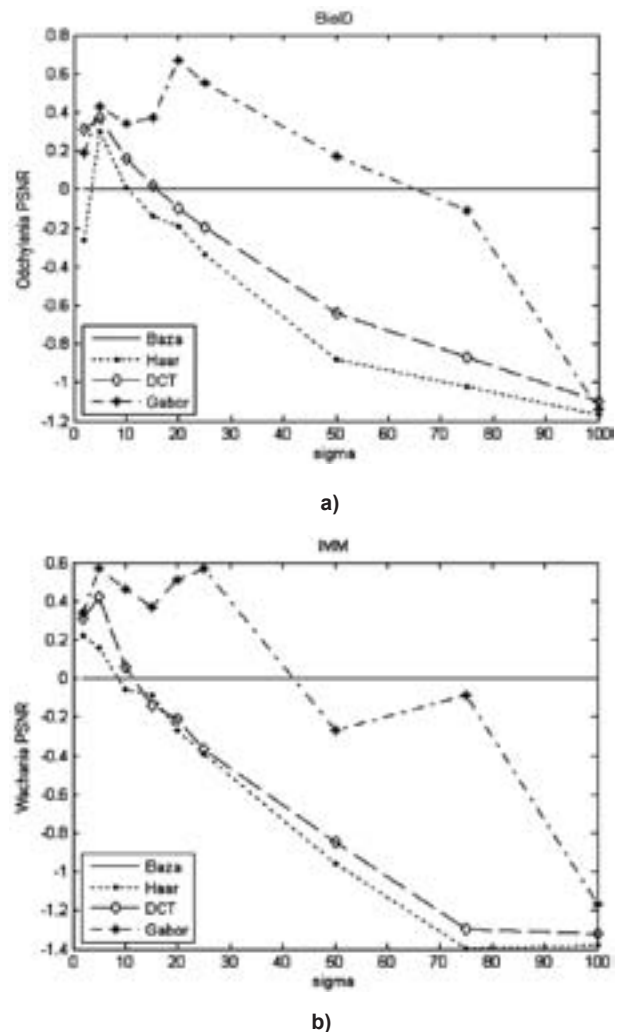


Rys. 8. Przykładowe rezultaty usuwania szumu do obrazu IMM 07-1: a) obraz zaszumiany $\sigma = 50$, PSNR = 14,15 dB, b) Haar PSNR = 30,83 dB, c) DCTPSNR = 30,94 dB, d) Gabor PSNR = 31,39 dB
Fig. 8. Example of the denoising results for the image IMM 07-1: a) noisy image $\sigma = 50$, PSNR = 14,15 dB, b) DCT PSNR = 29,81 dB, c) HaarPSNR = 30,21 dB, d) GaborPSNR = 31,74 dB

odszumianych obrazów dla $\sigma = 50$ (rys. 8) łatwo zauważyć że powstałe zniekształcenia były efektem rekonstrukcji obrazów z mało licznych zbioru atomów.

Porównanie prezentowanej metody z wynikami zawartymi w pracy [16] przedstawiono na rys. 9. Obrazują one fluktuacje jakości opisywanej metody usuwania szumu wokół linii bazowej (wartości osiągnięte dla obrazów testowych na podstawie metody opisanej przez Portilla w [16]) dla tych samych wartości parametru σ . Analizując te wykresy można stwierdzić, że dodatnie odchylenie PSNR zależne jest od typu wykorzystanego słownika, zawartości kontekstowej odszumianych obrazów oraz poziomu parametru σ . Dla wszystkich obrazów testowych najlepsze wyniki PSNR osiągnięto dla słownika Gabora w zakresie parametru $\sigma \in (0,50)$ co obrazują dodatnie odchylenia PSNR. Najgorsze wyniki PSNR uzyskano dla słownika Haara już dla $\sigma \geq 10$. Wynikało to prawdopodobnie z charakterystyki zmienności (rys. 4) atomów tego słownika.

Przedstawiono wykresy charakteryzujące średnią jakość wyliczoną dla całego zbioru obrazów z baz IMM oraz BioIDz wykorzystaniem słowników Haara, DCT i Gabora dla parametru $\sigma \in (0,100)$ modelującego addytywny szum gaussowski.



Rys. 9. Porównanie jakości odszumiania obrazów testowych z rezultatami uzyskanymi w pracy [16]
Fig. 9. Comparison between the denoising quality of sample images with the results achieved in work [16]

W celu porównania otrzymanych wyników przyjęto za poziom odniesienia (linia bazowa w zerze) rezultaty osiągnięte z wykorzystaniem metody opisanej w pracy [16] dla tych samych parametrów σ . Widoczne wahania względem linii bazowej obrazują skuteczność prezentowanej metody usuwania szumów z obrazów testowych.

Opisywaną metodę usuwania szumu wykorzystującą adaptacyjną dekompozycję obrazów zaimplementowano przy użyciu środowiska MATLAB 7.5 z wykorzystaniem biblioteki Image Processing Toolbox.

Podsumowanie

W artykule zaprezentowano metodę usuwania szumu z obrazów twarzy wykorzystującą nadmiarowe reprezentacje słownikowe obrazów. Opisano adaptacyjną metodę dekompozycji obrazu opartą na zachłannym algorytmie MP (*ang. Matching Pursuit*), który w każdej iteracji wyszukuje i dopasowuje elementy danego słownika do struktury analizowanego obrazu. Przebadano trzy typy słowników zawierających atomy funkcji: Harra, DCT i Gabora. W wyniku przeprowadzenia eksperymentów stwierdzono, że prezentowana metoda usuwania szumu osiąga najlepsze wyniki dla słownika modelowanego funkcjami Gabora w szerokim zakresie parametru σ modelowanego addytywnego szumu impulsowego. Potwierdza to przypuszczenia że duża różnorodność atomów w wykorzystywanym słowniku Gabora powoduje ich lepsze dopasowanie do elementów strukturalnych obrazu w procesach adaptacyjnej dekompozycji.

Dalsze prace ukierunkowane są na przebadanie licznego zbioru słowników opisanego różnymi typami funkcji bazowych w celu szukania odpowiedzi na pytanie, jak charakter poszczególnych funkcji wpływa na osiągnięte wyniki odszumiania.

Literatura

- [1] Aharon M., Elad M., Bruckstein A.: K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions Signal Processing*, 54, 11, pp. 4311–4322, 2006.
- [2] Banham M.R., Brailan J.C.: A selective update approach to matching pursuits video coding. *IEEE Transactions on Circuits and Systems Video Technology*, vol. 7, no. 1, pp. 119–129, 1997.
- [3] Bergeaud F., Mallat S.: Matching pursuit of images. In *Proc. IEEE International Conference on Image Processing ICIP'95*, vol. 1, pages 53–56, 1995.
- [4] Białasiewicz J.T.: Falki i aproksymacje. WNT Warszawa 2004.
- [5] Bruckstein A. M., Donoho D. L., Elad M.: From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.
- [6] Davis G., Mallat S., Avellaneda M.: Adaptive greedy approximations. *Journal of Constructive Approximation*, vol. 13, pp. 57–98, 1997.
- [7] Durka P.J.: Między czasem a częstotliwością: elementy współczesnej analizy sygnałów. 2004. <http://brain.fuw.edu.pl/~durka/as/>
- [8] Durka P.J.: Time-frequency analyses of EEG, dissertation. Institute of Experimental Physic, Warsaw University, 1996.
- [9] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, Springer, 2010.
- [10] Elad M., Aharon M.: Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions Image Processing*, vol. 54, pp. 3736–3745, 2006.
- [11] Gabor D.: Theory of communication. *Journal of Institution Electrical Engineering*, vol. 93, no. 26, pp. 429–457, 1946.
- [12] Haar A.: Zur Theorie der orthogonalen Funktionensysteme. *Mathematische Annalen*, vol. 69, pp. 331–371, 1910.
- [13] Janssen A.: Gabor representation of generalized functions, *Journal of the Mathematical. Analysis. and Applications*, vol. 83, no. 2, pp. 377–394, 1981.
- [14] Liu Q., Wang Q., Wu L.: Size of the dictionary in matching pursuit algorithm. *IEEE Transactions on Signal Processing*, vol. 52, no. 12, pp. 3403–3408, 2004.
- [15] Mallat S., Zhang Z.: Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.
- [16] Portilla J., Strela V., Wainwright M.J., Simoncelli E.P.: Image denoising using scale mixture of Gaussians in the wavelet domain. *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1338–1351, 2003.
- [17] Rubinstein R., Bruckstein M., Elad M.: Dictionaris for Sparse Representation Modeling. *Proceedings of the IEEE*, vol. 98, pp. 1045–1057, 2010.
- [18] Stegmann M., Ersboll B.K., Larsen R.: FAME – a flexible appearance modeling environment. *IEEE T. on Medical Imaging*, vol. 22, no. 10, 1319–1331, 2003.
- [19] Tropp J. A.: Greed is good: Algorithmic results for sparse approximation. ICES Report 03-04, The University of Texas at Austin, 2003.
- [20] <http://www.bioid.com/support/downloads/software/bioid-face-database.html>.