

Odszumianie obrazów

Seminarium: Algorytmy numeryczne i graficzne

Karolina Jeziorska, Oskar Bujacz

Wrocław, 8 czerwca 2020 r.

Plan prezentacji

- 1 Wstęp
- 2 Rodzaje szumów
- 3 Oceny jakości
- 4 Filtry ogólnie
- 5 Filtry w dziedzinie przestrzennej
- 6 Filtry w dziedzinie częstotliwości
- 7 Aktualne trendy i zastosowania

- sygnały nie są idealne
- szum i jego przyczyny
- cel: usunięcie szumu bez utraty jakości zdjęcia

Przedstawienie obrazu

Definicja

Wartość (jasność) obrazu w punkcie (x, y) określa funkcja f zadana wzorem:

$$f(x, y) = t(x, y) + s(x, y) \quad (1)$$

Gdzie funkcja t to prawdziwa wartość obrazu (sygnału), a s to dodany szum. Taki szum nazywamy szumem niezależnym.

Obraz wyjściowy



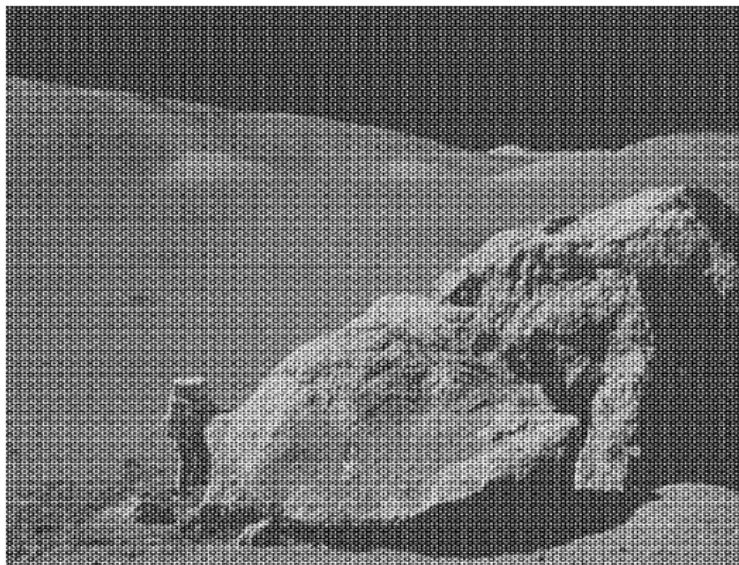
Szum sól i pieprz



Szum Gaussa



Szum okresowy



źródło: <https://hal.inria.fr/hal-01206546/file/ScipyLectures-simple.pdf>

RMSE

Definicja

Dla obrazu o rozmiarze $M \times N$ RMSE (pierwiastek błędu średniokwadratowego) definiujemy wzorem:

$$RMSE = \sqrt{\frac{1}{MN} \sum_{i=x}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - t(x, y)]^2} \quad (2)$$

SNR

Definicja

Dla obrazu o rozmiarze $M \times N$ SNR (stosunek sygnału do szumu) definiujemy wzorem:

$$SNR = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y)]^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - t(x, y)]^2} \quad (3)$$

PSNR

Definicja

Dla obrazu o rozmiarze $M \times N$ PSNR (szczytowy stosunek sygnału do szumu) definiujemy wzorem:

$$PSNR = \frac{[\max_{x,y} f(x, y)]^2}{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - t(x, y)]^2} \quad (4)$$

Decybel

SNR i PSNR często podawane są w decybelach, które są logarytmiczną skalą porównawczą, czyli dokładnie mówią o ile coś jest większe, mocniejsze od jakiegoś stanu zerowego.

Definicja

Wartość wyrażona w decybelach to iloraz danej wartości R do pewnej wartości odniesienia R_0 :

$$R_{dB} = 10 \log_{10} \frac{R}{R_0} \quad (5)$$

Filtre cyfrowe to algorytmy, które przekształcają sygnał cyfrowy (dyskretny) zgodnie z funkcją przejścia, która może być liniowa lub nie. Filtry są używane by zredukować lub wzmacnić pewne aspekty sygnału. Możemy podzielić je na dwa rodzaje:

- **Filtre dolnoprzepustowy (low pass)** przepuszczają sygnały poniżej ustalonej częstotliwości, a tłumią te o wyższej. Stosowane w redukcji szumów.
- **Filtre górnoprzepustowy (high pass)** przepuszczają sygnały powyżej ustalonej częstotliwości, a tłumią te o niższej. Stosowane do wyostrzania obrazów.

Definicja

Splot (konwolucja) to działanie określone dla dwóch funkcji, które daje w wyniku inną wyrażającą jak kształt jednej jest modyfikowany przez drugą. Działanie to określa się wzorami:

$$h(x) = (f * g)(x) = \int_{-\infty}^{+\infty} f(x - \tau)g(\tau)d\tau \quad (6)$$

$$h[x] = (f * g)[x] = \sum_{n=-\infty}^{+\infty} f[n]g[x - n] \quad (7)$$

Dla funkcji dwóch zmiennych w dziedzinie dyskretnej splot wyraża się wzorem:

$$h[x, y] = (f * g)[x, y] = \sum_{m,n} w[m, n]j[x - m, y - n] \quad (8)$$

Twierdzenie o konwolucji dla transformat Fouriera

Niech f i g będą dwoma funkcjami ze splotem $f * g$, a $\mathcal{F}\{f\}$ i $\mathcal{F}\{g\}$ to transformaty Fouriera tych funkcji. Wtedy zachodzi:

$$\mathcal{F}\{f * g\} = \mathcal{F}\{f\} \cdot \mathcal{F}\{g\} \quad (9)$$

$$f * g = \mathcal{F}^{-1}\{\mathcal{F}\{f\} \cdot \mathcal{F}\{g\}\} \quad (10)$$

Definicja

Filtrowanie liniowe obrazu f maską filtrującą w o wymiarze $m \times n$ wyraża się wzorem:

$$g(x, y) = \sum_{p=-a}^a \sum_{q=-b}^b f(x-p, y-q)w(p, q) \quad (11)$$

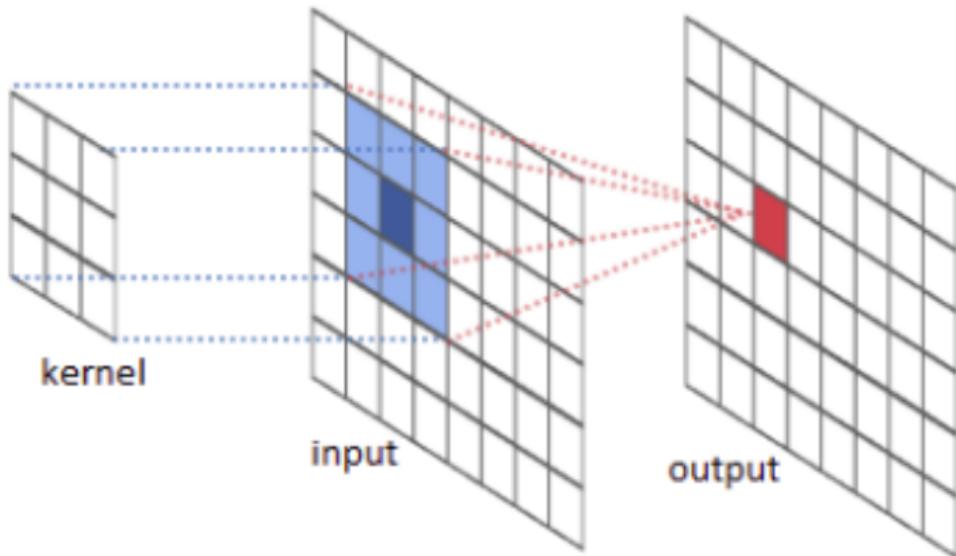
Gdzie $a = (m - 1)/2$, a $b = (n - 1)/2$. Można uprościć zapis i powiedzieć, że dla każdego punktu R obrazu jego wartość to:

$$R = w_1z_1 + w_2z_2 + \cdots + w_{mn}z_{mn} = \sum_{i=1}^{mn} w_i z_i \quad (12)$$

Gdzie w_i to współczynniki maski, a z_i to wartości odpowiadających im pikseli na wybranym obszarze.

Należy też pamiętać, że współczynniki maski powinny sumować się do 1, by uniknąć rozjaśnienia lub przyciemnienia obrazu. Jeżeli nie sumują się do 1, należy je pojedynczo przeskalać lub też przeskalać finalny wynik. Wtedy podany wcześniej wzór zmienia się na:

$$g(x, y) = \frac{\sum_{p=-a}^a \sum_{q=-b}^b f(x - p, y - q) w(p, q)}{\sum_{p=-a}^a \sum_{q=-b}^b w(p, q)} \quad (13)$$



źródło: <https://colah.github.io/posts/2014-07-Understanding-Convolutions/>

Średnia

Filtr uśredniający zmniejsza różnice w intensywności między pikselami poprzez zastąpienie wartość punktu wartością średniej jego sąsiadów (w tym też jego samego). Maska (jądro) wyznacza kształt i rozmiar sąsiedztwa, z którego wyliczamy średnią. Dla maski rozmiaru $m \times n$ wyznaczamy jej wartości jako:

$$w(p, q) = \frac{1}{mn} \quad (14)$$

Dla ułatwienia i przyspieszenia obliczeń możemy zmienić wartości maski na:

$$w(p, q) = 1 \quad (15)$$

Należy oczywiście wtedy pamiętać o finalnym przeskalowaniu wyniku.

Średnia

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Przykładowa maska uśredniająca rozmiaru 3×3

1	1	1
1	1	1
1	1	1

Zmieniona przykładowa maska uśredniająca rozmiaru 3×3

Średnia - wady

- piksel z wyróżniającą się wartością może mocno wpływać na wartość średniej
- wygładzenie (rozmycie) krawędzi

Średnia - modyfikacje

Aby zmniejszyć negatywne skutki zwykłego uśredniania, stosujemy średnią ważoną, gdzie największą wagę ma środkowy punkt (dla którego wartość liczymy). W ten sposób zmniejszamy wpływ otoczenia na jego wartość.

1	1	1
1	a	1
1	1	1

Przykładowa maska dla średniej ważonej rozmiaru 3×3

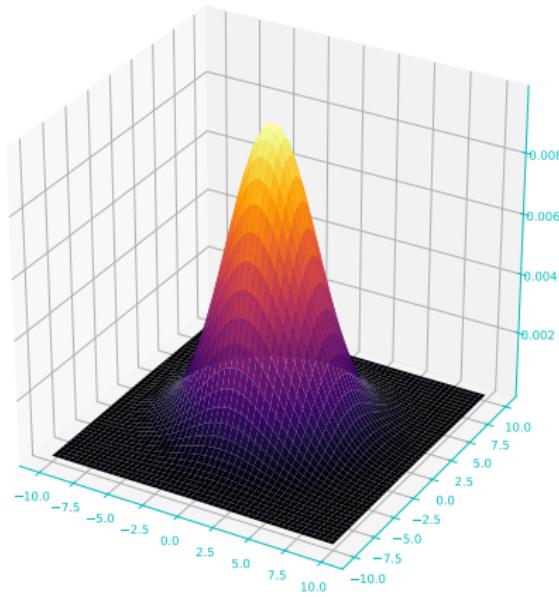
Gauss

Filtr Gaussa jest w pewnym rodzaju odmianą filtru średniej ważonej. W tym filtrze maska przyjmuje wartości rozkładu Gaussa wyśrodkowanego w środku maski: Oczywiście wagи sąsiadów mogą być różne. Często, by te wartości wyznaczyć, używa się rozkładu Gaussa:

$$w(p, q) = \frac{1}{2\pi\sigma^2} e^{-\frac{p^2+q^2}{2\sigma^2}} \quad (16)$$

W teorii rozkład Gaussa ma wszędzie wartości niezerowe (czyli maska powinna być nieskończona), jednak w rzeczywistości wartości odległe o więcej niż 3σ od środka przyjmują wartości bliskie 0, dlatego maska powinna mieć wymiary $6\sigma \times 6\sigma$

Dwuwymiarowy rozkład Gaussa



Gauss

Przykładowy filtr Gaussa, $\sigma = 0.5$

0.00719253	0.05314604	0.00719253
0.05314604	0.3926991	0.05314604
0.00719253	0.05314604	0.00719253

Gauss - заlety

- wartość σ reguluje poziom wygładzenia
- słabsze rozmycie krawędzi niż zwykły filtr średniej
- rozdzielność

$$\frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} \quad (17)$$

Definicja

Filtrowanie lokalne nieliniowe wyraża się wzorem:

$$g(x, y) = \phi(f(x, y)) \quad (18)$$

Gdzie zbiorem wartości funkcji ϕ są wartości pikseli z otoczenia punktu (x, y) .

Median

Filtr mediany eliminuje te piksele, które w znacznym stopniu odbiegają od pozostałych pikseli. Dokładnie filtr wyraża się wzorem:

$$g(x, y) = \text{median}_{p,q \in K}(f(x + p, y + q)) \quad (19)$$

Gdzie K to maska (okno) filtra, czyli piksele sąsiadujące z (x, y) .
Medianą dużo lepiej zachowuje krawędzie i jest bardzo skuteczna przy usuwaniu szumu typu sól i pieprz.

Mediana - wady

- eliminacja faktury, stworzenie obszarów o stałej wartości
- usunięcie detali o rozmiarze proporcjonalnym do rozmiaru maski
- obcinanie rogów
- długość obliczeń

Konserwatywne wygładzanie

Konserwatywne wygładzanie zapewnia, że wartość każdego piksela znajduje się w ograniczonym przez jego sąsiadów zakresie.

Niech $\min = \min_{p,q \in K} f(x+p, y+q)$ i $\max = \max_{p,q \in K} f(x+p, y+q)$ z pominięciem $p = q = 0$. Wtedy filtr wyraża się wzorem:

$$g(x, y) = \begin{cases} \min & f(x, y) < \min \\ \max & f(x, y) > \max \\ f(x, y) & \text{wpp} \end{cases} \quad (20)$$

Wynik konserwatywnego wygładzania jest dużo bardziej subtelny niż zwykłej mediany czy średniej.

Strojony filtr medianowy

Strojony filtr medianowy то поєднання середньої і медіан. Понownie маска означає залась значень. Усувають з них $d/2$ найменших і $d/2$ найбільших значень. Втіди фільтр виражається формулою:

$$g(x, y) = \frac{1}{n^2 - d} \sum_{p, q \in K} f(x + p, y + q) \quad (21)$$

где $d = 0, \dots, n^2 - 1$.

Для $d = 0$ отримуємо середню, $d = n^2 - 1$ медіану, а для решти значень фільтри посередині, які краще справляються з сумом.

Nielokalna średnia

W przeciwieństwie do średniej lokalnej, żeby określić wartość danego piksela, bierzemy średnią ze wszystkich pikseli na obrazku. Jednak wagi tych pikseli są zależne od tego, jak podobne są do punktu (x, y) . Czystość obrazu, jak i zachowanie krawędzi jest dużo lepsze niż w zwykłej średniej.

Nielokalna średnia

Algorytm

Algorytm nielokalnej średniej dla obrazu o wymiarach $M \times N$ wyraża się wzorem:

$$g(x, y) = \frac{1}{C(x, y)} \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} f(p, q) W(x, y, p, q) \quad (22)$$

Gdzie $C(x, y)$ to wartość skalująca:

$$C(x, y) = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} W(x, y, p, q) \quad (23)$$

Nielokalna średnia

Algorytm

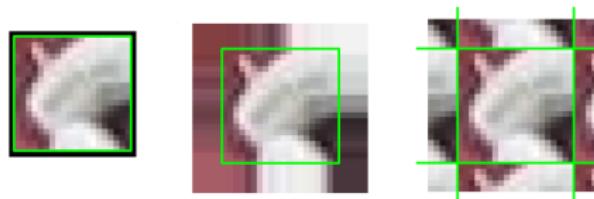
Funkcja wagowa W określa podobieństwo dwóch punktów na podstawie ich otoczenia. Stosuje się różne funkcje wagowe, przykładowo może to być funkcja wagowa Gaussa:

$$W(x, y, p, q) = e^{-\frac{(B(x,y)-B(p,q))^2}{h^2}} \quad (24)$$

Gdzie h to parametr filtru, a funkcja B zwraca lokalną średnią punktu (czyli średnią z jego otoczenia).

Problem brzegów

- ograniczenie obszaru działania maski
- filtrowanie pikseli brzegowych tylko częścią maski
- zwiększenie obszaru wejściowego poprzez dodanie *paddingu*



źródło: <http://aragorn.pb.bialystok.pl/~boldak/>

Czy możemy zmienić dziedzinę?

W dziedzinie przestrzennej pojawiają się problemy, które trudno przeskoczyć

- rozmywanie obrazu
- problem z wykrywaniem krawędzi
- efekt schodków np. przy ADF

Rozwiążanie - zmiana dziedziny, w której wykonujemy filtrowanie. Jedną z pierwszą propozycji była transformacja Fouriera, ale od tamtej pory wykonano wiele innych prób

- DCT
- falki
- sterowane falki
- wedgelets
- itd

Opis metody

Filtры в дзедзине частотлівости в порównaniu do фільтров пространственных, не выполняют ошумления непосредственно на изображении. Сперва они выполняют определенную трансформацию в частотную область и на основе ее свойств очищают изображение. Позже, конечно, возвращаются в пространственную область.

Wprowadzenie FT

Jednym z pomysłów jest transformacją Fouriera, którą wszyscy już świetnie znamy :). Przypomnimy tylko najważniejsze informacje.
FT w 1D definiujemy jako

$$F(u) = \int_{-\infty}^{\infty} f(x)e^{-i2\pi ux} dx \quad (25)$$

gdzie $i^2 = -1$ Oczywiście znamy też odwrotną FT

$$f(x) = \int_{-\infty}^{\infty} F(u)e^{i2\pi ux} du \quad (26)$$

Te dwa równania tworzą *Fourier transform pair*, czyli po zastosowanie jednej i drugiej wracamy do oryginalnego obrazu.

FT 2D

Naturalnie, możemy rozszerzyć rozumowanie na 2D

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-i2\pi(ux+vy)} dx dy \quad (27)$$

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{i2\pi(ux+vy)} du dv \quad (28)$$

Przedstawione wzory dotyczą funkcji ciągłych, ale w temacie obróbki obrazów wystarczą nam ich uproszczone wersje - transformacje dyskretnie (**DFT**).

DFT

DFT 1D

Definiujemy ją następująco

$$F(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(x) e^{-i2\pi ux/M} \quad \text{dla } u = 1, \dots, M-1 \quad (29)$$

$$f(x) = \sum_{u=0}^{M-1} F(u) e^{i2\pi ux/M} \quad \text{dla } x = 1, \dots, M-1 \quad (30)$$

Co ważne nie musimy się przejmować istnieniem DFT czy też odwrotności, ponieważ zakładamy, że wszystkie wartości pikseli są skończone.

DFT 2D

DFT 2D

Najważniejsze dzisiaj wzory. DFT rozszerzone na 2D

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-i2\pi(ux/M+vy/N)} \quad (31)$$

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{i2\pi(ux/M+vy/N)} \quad (32)$$

Warto powiedzieć, że w implementacji będziemy stosować FFT, aby poprawić złożoność z $O(n^2)$ do $O(n \log n)$.

Odszumianie obrazów

Filtры в дзидине częstotliwości

Dyskretna transformacja Fouriera

Lena standardowo

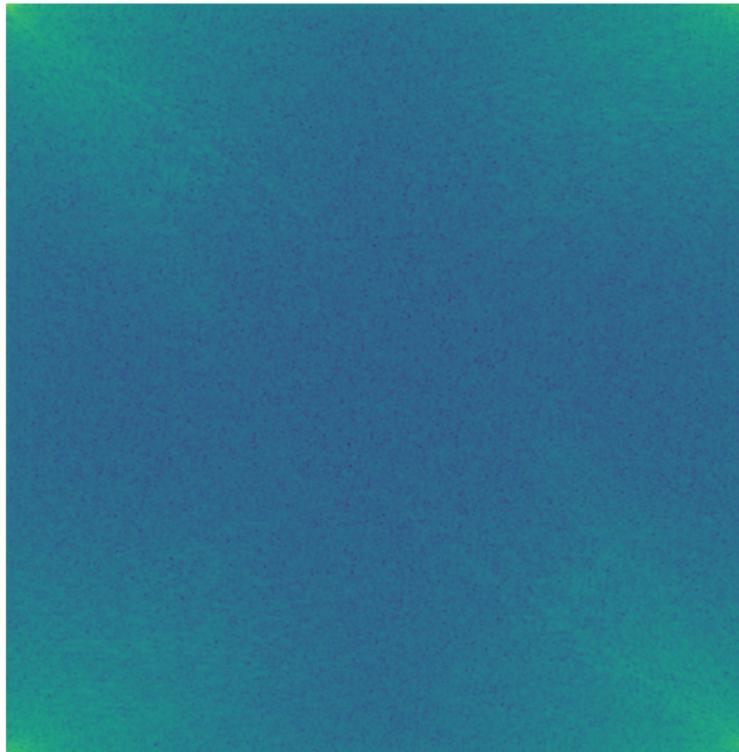


Odszumianie obrazów

Filtры в дзидзине częstotliwości

Dyskretna transformacja Fouriera

Lena w częstotliwości

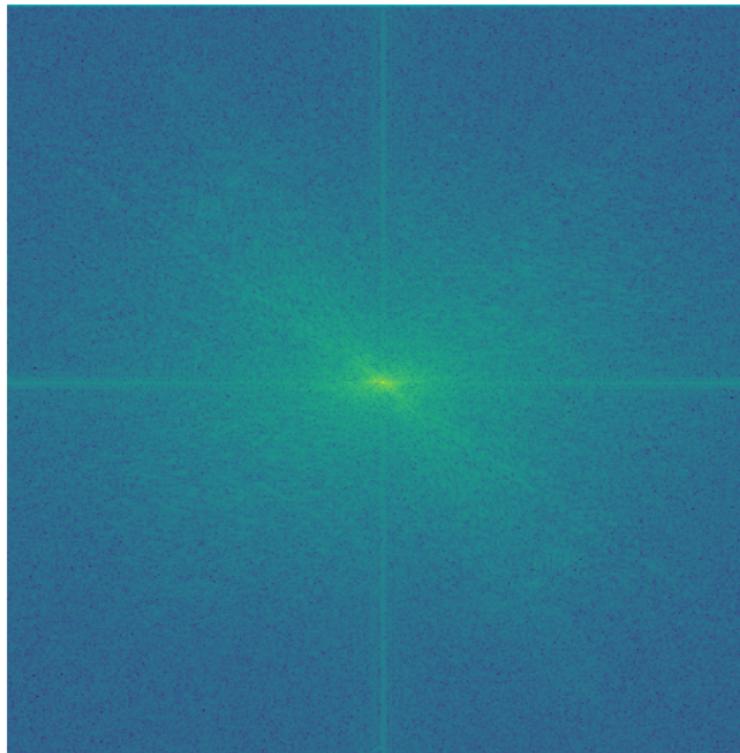


Odszumianie obrazów

Filtры в дзидзине częstotliwości

Dyskretna transformacja Fouriera

Lena w częstotliwości



Ważnym aspektem filtrowania w dziedzinie częstotliwości jest nałożenie przesunięcia S , które pozwala umieścić punkt odpowiadający ($u = v = 0$) na środku obrazu w punkcie ($u/2, v/2$). Tą operację wykonujemy następująco

- podział obrazu na 4 ćwiartki
- zamiana naprzeciwległych ćwiartek

Ułatwia to intuicyjne zrozumienie rozkładu na składowe częstotliwości oraz nakładanie filtrów. Wzór na punkt środkowy upraszcza się wtedy do

$$F(0, 0) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \quad (33)$$

czyli po prostu średnia intensywności wszystkich pikseli.

Wydawałoby się, że nie możemy zatem wyciągnąć wniosków o obrazie po zastosowaniu transformacji. Wiemy, że częstotliwość jest bezpośrednio powiązana z tempem zmian, zatem intuicyjnie możemy powiązać częstotliwości ze wzorami zmian w obrazie.

- blisko środka - gładkie obszary
- daleko od środka - detale, krawędzie, szum

Kroki filtrowania

- ① obliczamy $F(u, v)$ przez DFT
- ② wykonujemy przekształcenie S w celu wyśrodkowania transformacji
- ③ mnożymy $F(u, v)$ przez filtr $H(u, v)$
- ④ ponownie wykonujemy S
- ⑤ obliczamy odwrotną DFT
- ⑥ zostawiamy tylko części rzeczywiste

Funkcję $H(u, v)$ nazywamy filtrem, ponieważ obniża one pewne częstotliwości, pozostawiając inne niezmienione.

Idealny filtr dolnoprzepustowy

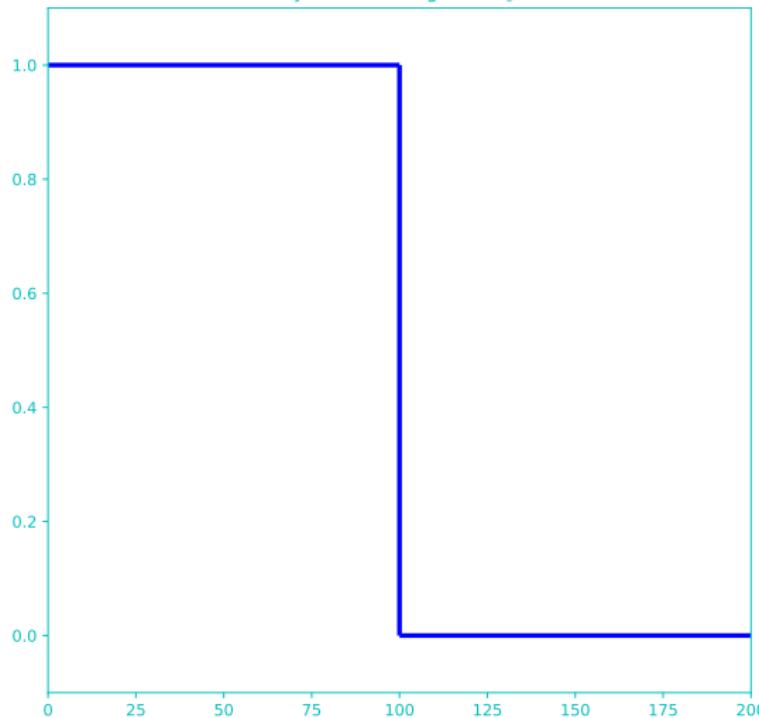
Будемо розглядати 3 типи фільтрів, першим з них буде ідеальний фільтр долнoprzepustowy. Визначимо його як

IFDP

$$H(u, v) = \begin{cases} 1 & D(u, v) < D_0 \\ 0 & D(u, v) > D_0 \end{cases} \quad (34)$$

где

$$D(u, v) = \sqrt{(u - M/2)^2 + (v - N/2)^2} \quad (35)$$

Przekrój filtra idealnego dla $D_0 = 100$ 

Widmowa gęstość mocy

Określa ona dystrybucję energii w poszczególnych składowych częstotliwościach naszego sygnału. Definiujemy ją jako

$$P(u) = |F(u)|^2 = R^2(u) + I^2(u) \quad (36)$$

gdzie R i I to odpowiednio rzeczywista i urojona część liczby zespolonej. Jest ona przydatna przy ocenie jakości odszumiania i procesie wybory progu dla filtrowania.

Ocena ilości energii w obrazie

Jak ocenić D_0 poza metodą prób i błędów? Szukamy procenta mocy.

Procent mocy na danym promieniu

$$P_T = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} P(u, v), \quad \alpha = 100 \sum_u \sum_v P(u, v) / P_T \quad (37)$$

Na podstawie tego wskaźnika jesteśmy w stanie stwierdzić, jak dużo widmowej gęstości energii pozostało po zastosowaniu filtra, co można rozumieć jako ilość zachowanej informacji.

Filtr dolnoprzepustowy Butterworth'a

Definiujemy go jako

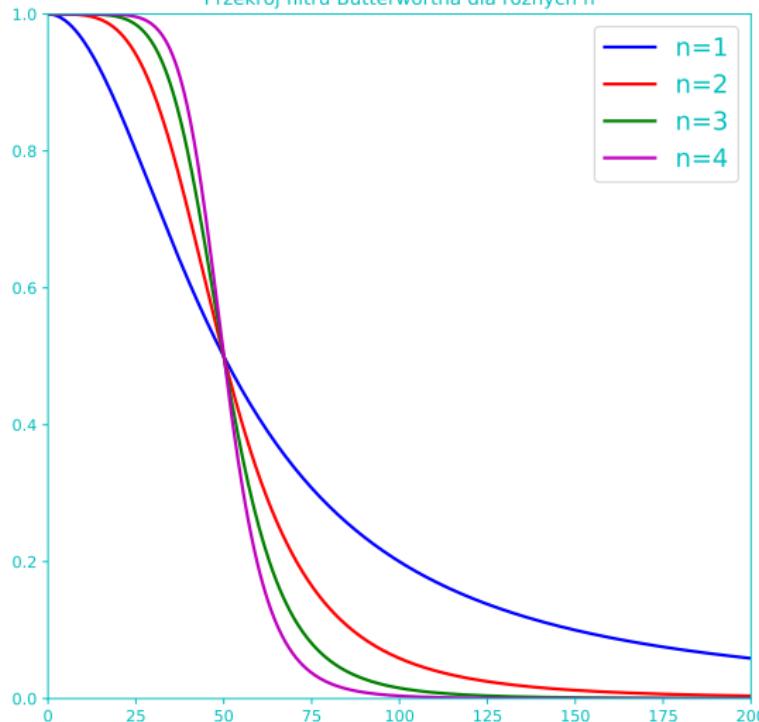
FDPB

$$H(u, v) = \frac{1}{1 + (D(u, v)/D_0)^{2n}} \quad (38)$$

gdzie n to parametr filtra.

Parametr pozwala nam modyfikować jak gwałtowna bądź łagodna jest zmiana akceptowanych składowych. Zaletą jest możliwość wyeliminowania zjawiska ringingu przy odpowiednim n .

Przekrój filtru Butterwortha dla różnych n



Gaussowski filtr dolnoprzepustowy

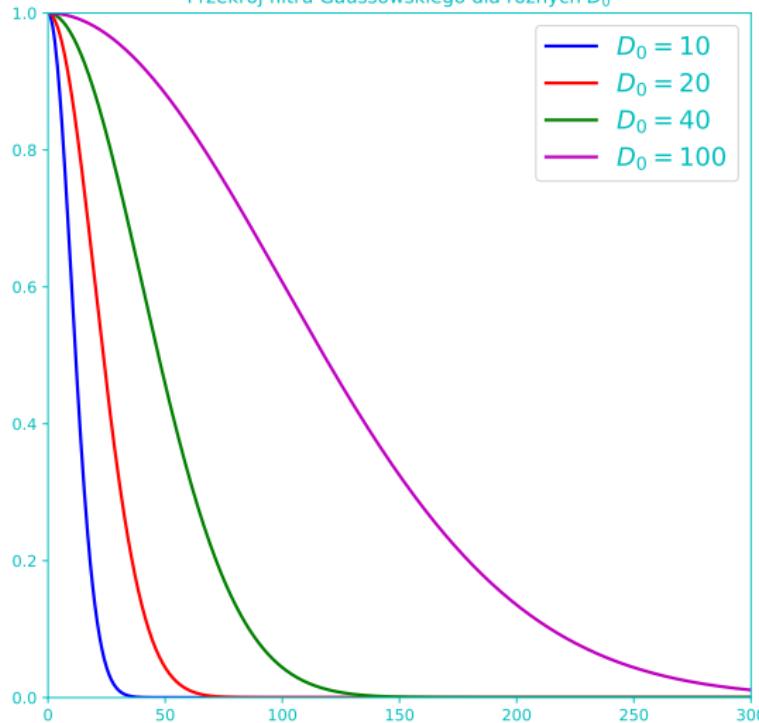
GFDP

$$H(u, v) = e^{-D^2(u, v)/2\sigma^2} \quad (39)$$

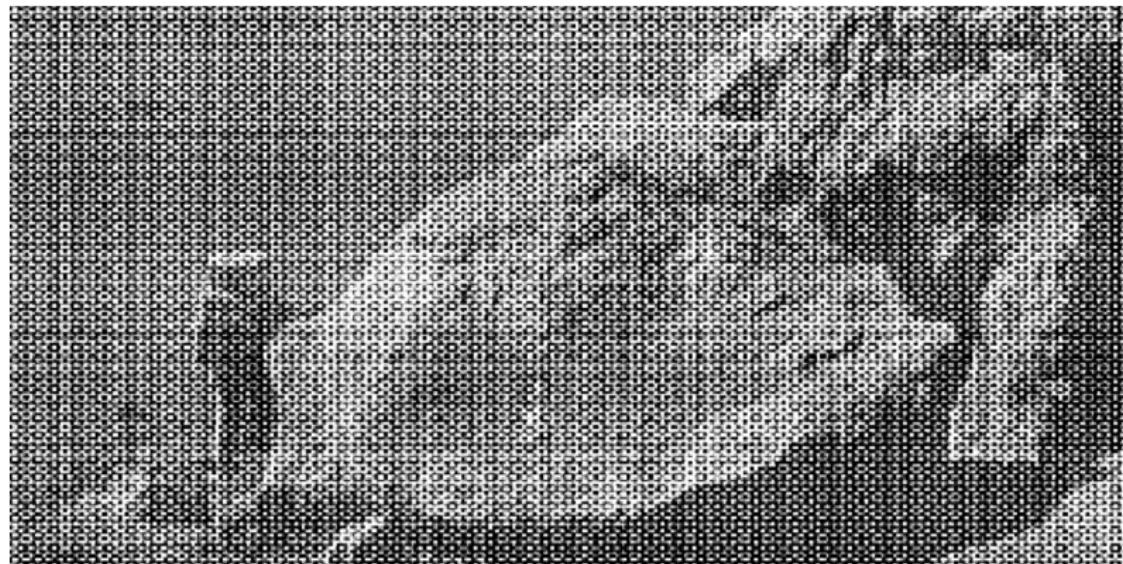
σ jest parametrem, który mówi nam o tym, jak szybko osłabiamy kolejne częstotliwości wraz ze zwiększeniem się składowych.

Będziemy stosować $\sigma = D_0$, wtedy wzór ma postać

$$H(u, v) = e^{-D^2(u, v)/2D_0^2} \quad (40)$$

Przekrój filtra Gaussowskiego dla różnych D_0 

Przykład - moonlanding

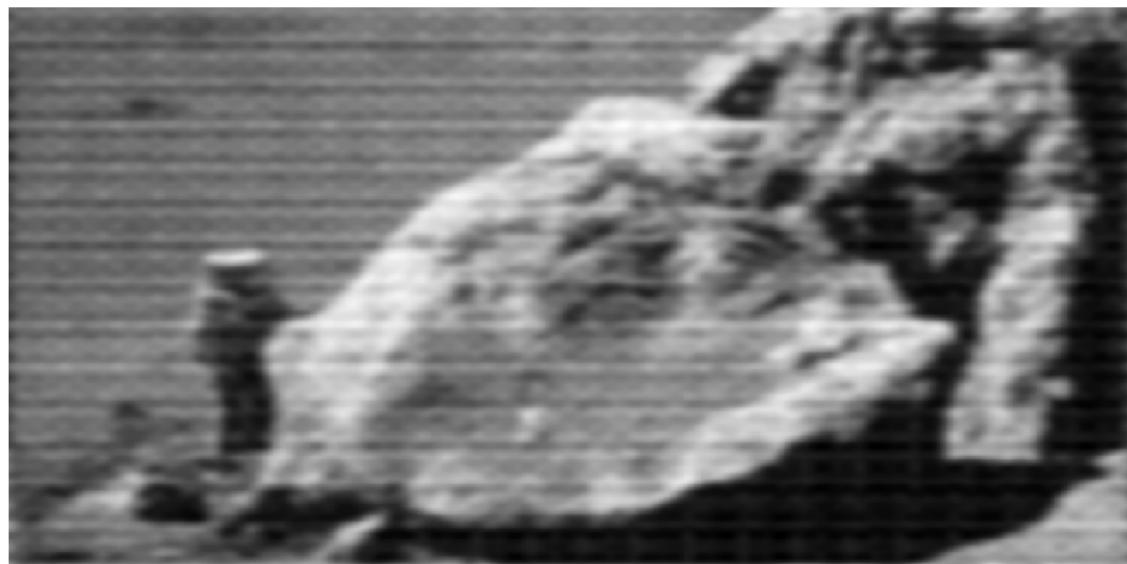


źródło: <https://hal.inria.fr/hal-01206546/file/ScipyLectures-simple.pdf>

Odszumianie obrazów

Filtr w dziedzinie częstotliwości

Filtryle



Dzisiejsze metody

Od prostych filtrów lokalnych, poszukiwania podobieństw w innych obszarach w filtrach nielokalnych, próby wykorzystania gradientów do wykrywania krawędzi, przez transformację częstotliwości, aż do głębokich sieci neuronowych - bardzo trudno jest opisać wszystkie metody usuwające szum z obrazów.

- użycie rzadziej i nadmiarowej reprezentacji na uczonych słownikach
- transformacja falkowa
- BM3D - block-matching and 3D filtering
- sieci neuronowych
- metody hybrydowe

Porównanie metod



Lena - szum dla $\sigma = 50$



Lena - odszumianie FFT



Lena - odszumianie BM3D

Zastosowania

- medyczne - zdjęcia rentgenowskie, tomografia komputerowa
- zdjęcia podwodne
- radary z syntetyczną aperturą
- zdjęcia w podczerwieni
- smartfony

Koniec

Bibliografia I

- [1] Witold Malina, Maciej Smiatacz *Cyfrowe przetwarzanie obrazów*. Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2008
- [2] Cezary Bołdak *Wykład: cyfrowe przetwarzanie obrazów* Politechnika Białostocka
- [3] Bhawna Goyal, Ayush Dogra, Sunil Agrawal, B.S. Sohi, Apoorav Sharma *Image denoising review: From classical to state-of-the-art approaches* Information Fusion 55, 2020
- [4] Antoni Buades, Bartomeu Coll, Jean-Michel Morel *A review of image denoising algorithms, with a new one* Multiscale Modeling Simulation Volume 4 Issue 2, 2005
- [5] Rafael C. Gonzalez, Richard E. Woods *Digital Image Processing* Prentice Hall 2002

Bibliografia II

- [6] <http://homepages.inf.ed.ac.uk/rbf/HIPR2/>