



Spis treści
Wstęp
Sygnał jednowymiarowy
Sygnał dwuwymiarowy
Obraz kolorowy



Algorytmy przetwarzania sygnałów i obrazów

Filip Piękniewski

Wydział Matematyki i Informatyki
UMK

Prezentacja dostępna na
<http://www.mat.uni.torun.pl/~philip/>

3 maja 2008



- 1 Wstęp
- 2 Sygnał jednowymiarowy
 - Podstawowe cechy sygnału i próbkowania
 - FFT jako ewaluacja wielomianu
 - Twierdzenie o próbkowaniu
 - Podstawowe filtry splotowe - idea
- 3 Sygnał dwuwymiarowy
 - FFT w dwóch wymiarach
 - Sploty dwuwymiarowe
 - Rozplatanie (deconvolution)
 - Transformata Radona i tomografia
 - Eliminacja szumu punktowego
 - Aliasing w 2d
- 4 Obraz kolorowy
 - Kolory i percepcja
 - Reprezentacje kolorów
 - Zastosowania

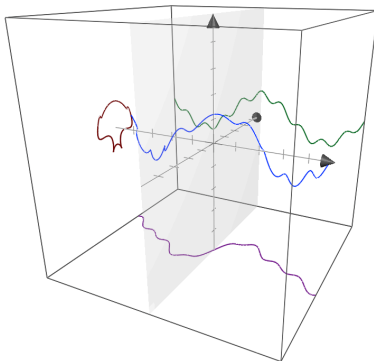


Przetwarzanie sygnałów to ważny kawałek szeroko pojętej informatyki. Specyfika tej dziedziny wymaga pewnego przygotowania fizycznego i jest czasem trudno zrozumiała dla ludzi zajmujących się bardziej czystą informatyką". Podstawowym pojęciem tej dziedziny jest transformata Fouriera, która jest obecna w programie praktycznie każdego studiów informatycznych, jednak jej elegancja rzadko bywa doceniona przez studentów. W dalszej części prezentacji zawartych jest sporo przykładów z życia wziętych, także w przypadku obrazów (sygnałów dwuwymiarowych). Na koniec jest trochę na temat reprezentacji kolorów.

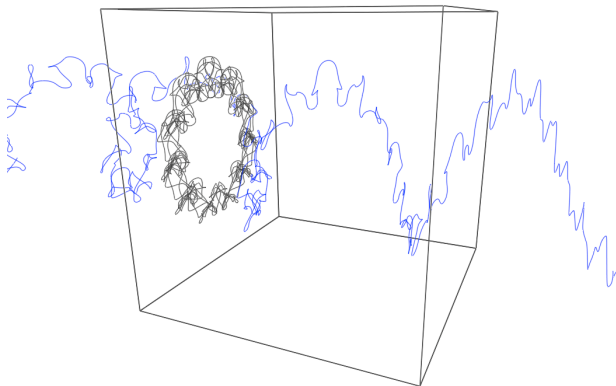


Definicja

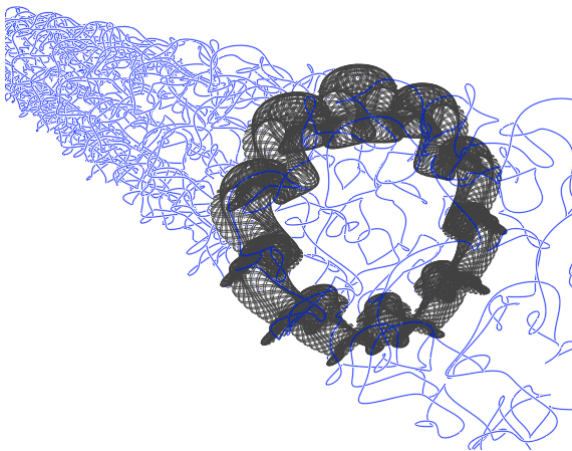
Sygnałem jednowymiarowym będziemy nazywali dowolną funkcję $f : \mathbb{R} \rightarrow \mathbb{C}$ należącą do przestrzeni L^1 . Ponadto gdy f przyjmuje jedynie wartości rzeczywiste mówimy o sygnale rzeczywistym. Jeśli f należy do L^2 mówimy o sygnale o skończonej energii. Sygnały zespolone cechuje polaryzacja - kształt rzutu sygnału na płaszczyznę zespoloną. Najczęściej spotykane są polaryzacje liniowe, kołowe i eliptyczne. W ogólności polaryzacja może mieć kształt dowolnej krzywej Lissajous.



Rysunek: Wizualizacja sygnału zespolonego - oś idącą od lewej do prawej należy kojarzyć z upływem czasu, płaszczyzna prostopadła do osi czasu, to dziedzina sygnał ($\mathbb{C}$), rzut na płaszczyznę prostopadłą (cień) to charakterystyka polaryzacji sygnału.



Rysunek: Inny przykład sygnału zespolonego i bardziej skomplikowanej polaryzacji.



Rysunek: Inny przykład sygnału zespolonego i bardziej skomplikowanej polaryzacji.



W dalszym ciągu będą nas interesowały przekształcenia sygnału, zwane czasem filtrami, szczególnie takie:

- które są liniowe, zatem aplikacja filtru na sumie sygnałów to to samo co suma aplikacji na składowych
- filtr nie zależy od czasu, zatem aplikacja filtru dla sygnału przesuniętego w czasie o Δt daje ten sam wynik przesunięty o Δt .

Filtry takie są jednoznacznie charakteryzowane tak zwaną odpowiedzią impulsową (*impulse response*), czyli wynikiem aplikacji filtru na sygnale $\delta_{t_0}(t)$ (delta Diraca). Ponadto ze względów praktycznych rozważa się *filtry realizowalne* których wynik zależy jedynie od części sygnału znajdującego się w przedziale $(-\infty, t)$.



Założmy, że mamy pewien sygnał $a(t)$. Rozważmy wartości $a(t)$ jedynie dla pewnych wybranych T punktów oddalonych od siebie o 1. Taki okrojony sygnał $\tilde{a}(t)$ nazwiemy sygnałem spróbkowanym na przedziale $[0, T)$. Zauważmy, że:

$$\tilde{a}(t) = \sum_{i=0}^{T-1} a(i)\delta(t - i) \quad (1)$$

Ponadto założmy, że będziemy aplikować pewien filtr b , oraz znamy odpowiedź impulsową filtra $b(t)$. Jeśli b jest liniowy i inwariantny czasowo, wtedy odpowiedź filtra b na sygnał $\tilde{a}$ możemy łatwo wyliczyć:

$$c(k) = \sum_{i=0}^k \tilde{a}(i)b(k - i) \quad (2)$$



Przypatrzmy się jeszcze raz wzorowi:

$$c(k) = \sum_{i=0}^k \tilde{a}(i)b(k-i) \quad (3)$$

Zauważmy, że $c(k)$ to dokładnie współczynnik stojący przy k -tej potęgde w iloczynie wielomianów o współczynnikach $\tilde{a}(i)$ oraz $b(i)$!
Zatem efektywna aplikacja filtru jest równoważna szybkemu mnożeniu wielomianów!



Jak efektywnie mnożyć wielomiany? Ustalmy dla porządku:

$$A(x) = a_d x^d + a_{d-1} x^{d-1} + \dots + a_1 x + a_0 \quad (4)$$

Zauważmy, że każdy wielomian stopnia d jest jednoznacznie reprezentowany przez:

- zbiór $d + 1$ współczynników $a_d, a_{d-1}, \dots, a_1, a_0$
- zbiór wartości w $d + 1$ różnych punktach
 $A(x_0), A(x_1), \dots, A(x_d)$

W reprezentacji przez współczynniki, wymnożenie wielomianu wymaga dla każdego c_k , k kroków, co ostatecznie daje $\mathcal{O}(d^2)$.



Zauważmy ponadto, że iloczyn wielomianów $C(x) = A(x) * B(x)$ jest jednoznacznie charakteryzowany przez swoje wartości w przynajmniej $2d + 1$ punktach, oraz:

$$C(x_k) = A(x_k) * B(x_k) \quad (5)$$

Zatem w reprezentacji przez wartości mnożenie wielomianów zajmuje jedynie $2d + 1$ czyli $\mathcal{O}(d)$ czasu! Tyle, że zazwyczaj wielomiany są zadane przez współczynniki... Potrzebujemy zatem efektywnej procedury konwersji od współczynników do wartości (ewaluacji) oraz operacji odwrotnej (interpolacji). Niestety ewaluacja wielomianu w punkcie, zajmuje $\mathcal{O}(d)$ czasu, potrzebujemy $\mathcal{O}(d)$ punktów, więc znowu mamy $\mathcal{O}(d^2)$. Chyba że sprytnie dobierzemy punkty!



Zauważmy najpierw, że jeśli dobierzemy punkty tak, aby były parami dodatni-ujemny, to możemy oszczędzić trochę obliczeń. Popatrzmy na wielomian:

$$5x^5 + 4x^4 + 3x^3 + 2x^2 + x + 2 = (2 + 2x^2 + 4x^4) + x(5x^4 + 3x^2 + 1)$$

Jak widać składowe w nawiasach są wielomianami od x^2 . Ogólnie mamy:

$$A(x) = A_{\text{even}}(x^2) + xA_{\text{odd}}(x^2) \quad (6)$$

Wtedy

$$\begin{aligned} A(x_i) &= A_{\text{even}}(x_i^2) + xA_{\text{odd}}(x_i^2) \\ A(-x_i) &= A_{\text{even}}(x_i^2) - xA_{\text{odd}}(x_i^2) \end{aligned} \quad (7)$$

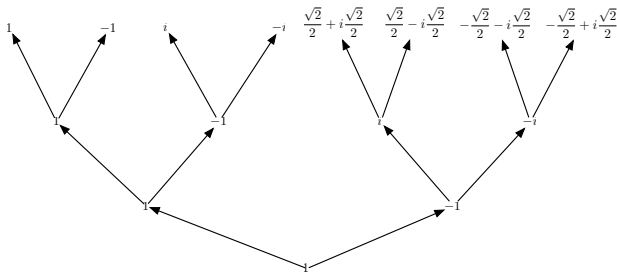
Ostatecznie zamiast ewaluować w n punktach, dwukrotnie ewaluujemy w $n/2$ punktach.



Gdyby dało się ten sam trick zastosować rekurencyjnie,
dostalibyśmy równanie rekurencyjne na złożoność:

$$T(n) = 2 * T(n/2) + \mathcal{O}(n) \quad (8)$$

co daje $T(n) = \mathcal{O}(n \log n)$. Ale to oznacza, że x_i^2 same muszą być parami dodatni-ujemny, a takich liczb rzeczywistych nie ma... I tutaj z pomocą przychodzą zespolone pierwiastki z jedynek...



Rysunek: Schemat rozgałęzień w FFT. Na n -tym poziomie są pierwiastki z jedynki stopnia 2^n



Pierwiastki n -tego stopnia z jedności $1, \omega, \omega^2, \dots, \omega^{n-1}$ wyrażają się łatwo wzorem $\omega = e^{2\pi i/n}$. Ponadto gdy n jest parzyste mamy:

- pierwiastki tworzą pary dodatni-ujemny bo $\omega^{n/2+j} = -\omega^j$
- podniesienie do kwadratu pierwiastków stopnia n daje pierwiastki stopnia $n/2$

Z tego względu liczby te doskonale nadają się do naszego rekurencyjnego algorytmu! Drobną niedogodnością jest fakt, iż wszystko działa doskonale o ile $n = 2^k$. Gdy jest inaczej, rekursję trzeba przerwać na pewnym poziomie i dokonać ewaluacji ręcznie... Można jednak ominąć tą niedogodność, przez uzupełnienie zerami do potęgi dwójki.



Mamy zatem procedurę, zwaną szybką transformatą Furiera (Fast Fourier Transform - FFT) która ewaluje wielomian w punktach będących pierwiastkami zespolonymi jedności. W tej reprezentacji możemy efektywnie mnożyć wielomiany, ale jak wrócić do reprezentacji współczynnikowej? Innymi słowy co z interpolacją?

Aby znaleźć odpowiedź przyjrzymy się jeszcze raz ewaluacji, tym razem w postaci macierzowej.

Ewaluacja wielomianu $A(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ w punktach $x_0, x_1, \dots, x_n$ wyraża się mnożeniem macierzy M (w środku):

$$\begin{bmatrix} A(x_0) \\ A(x_1) \\ \vdots \\ A(x_n) \end{bmatrix} = \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^{n-1} \\ 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \dots & x_{n-1}^{n-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}$$

Interpolacja zatem sprowadza się do odwrócenia macierzy M i wymnożenia wartości przez M^{-1}



W naszym przypadku macierz M ma dość specyficzną postać:

$$M_n(\omega) = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(n-1)} \\ & & \vdots & & \\ 1 & \omega^j & \omega^{2j} & \dots & \omega^{j(n-1)} \\ & & \vdots & & \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \dots & \omega^{(n-1)(n-1)} \end{bmatrix} \begin{matrix} \leftarrow \omega^0 = 1 \\ \leftarrow \omega \\ \leftarrow \omega^2 \\ \vdots \\ \leftarrow \omega^j \\ \vdots \\ \leftarrow \omega^{n-1} \end{matrix}$$

Jest to macierz Vandermonda którą można odwrócić w czasie $\mathcal{O}(n^2)$ (zamiast $\mathcal{O}(n^3)$) jednak to nadal zbyt wolno.



Zauważmy jeszcze jeden fakt - kolumny macierzy $M_n(\omega)$ są ortogonalne. Łatwo to sprawdzić licząc iloczyny skalarne ¹:

$$u \cdot v^* = u_0 v_0^* + u_1 v_1^* + \dots + u_{n-1} v_{n-1}^*$$

W przypadku iloczynu j -tej i k -tej kolumny $M_n(\omega)$ mamy:

$$1 + \omega^{j-k} + \omega^{2(j-k)} + \dots + \omega^{(n-1)(j-k)}$$

Jak widzimy jest to szereg geometryczny o ilorazie ω^{j-k} którego suma wynosi:

$$\frac{1 - \omega^{n(j-k)}}{1 - \omega^{j-k}} = \begin{cases} 0 & j \neq k \\ n & j = k \end{cases}$$

¹* oznacza sprzężenie zespolone



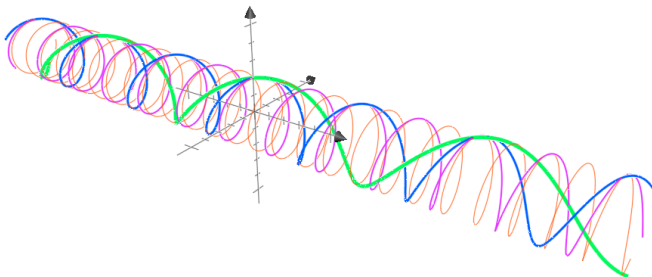
Poprzednią własność można zapisać równaniem macierzowym:

$$M_n(\omega)M_n(\omega)^* = nI \quad (9)$$

Gdzie I jest macierzą identycznościową. Ale $M_n(\omega)^* = M_n(-\omega)$. Przenosząc skalar na lewo i mnożąc lewostronnie przez $M_n(\omega)^{-1}$ dostajemy:

$$\frac{1}{n}M_n(-\omega) = M_n(\omega)^{-1} \quad (10)$$

Zatem operacją odwrotną do transformaty Fouriera jest sama transformata Fouriera uzupełniona o czynnik $1/n$ i wzięta ze sprzężeniem! To oznacza, że interpolacja też jest w czasie $\mathcal{O}(n \log n)$!



Rysunek: Kolumny macierzy $M_n(\omega)$ należy kojarzyć z punktami równo rozmieszczonymi na coraz bardziej ściśniętych sprężynach. Zbiór wszystkich takich sprężyn stanowi bazę ortogonalną całej przestrzeni Hilberta L^2 .



Twierdzenie o próbkowaniu

Twierdzenie (Nyquist-Shannon)

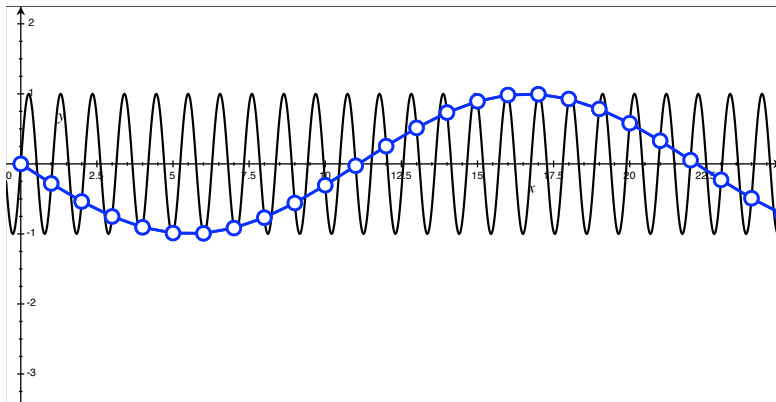
Jeśli funkcja f ma widmo zerowe powyżej częstotliwości $\omega_{\max}$, wtedy jest całkowicie zdeterminowana przez ciąg wartości w punktach odległych o $\frac{1}{2\omega_{\max}}$ (w dziedzinie czasu). Częstotliwość będącą połową częstotliwości próbkowania zwana jest częstotliwością Nyquista.

Twierdzenie to oznacza, iż sygnały o zwartym widmie, pomimo iż mogą być bardzo skomplikowane, są podobnie jak wielomiany determinowane przez skończony ciąg wartości w punktach (lub skończony ciąg współczynników widmowych) .



Twierdzenie o próbkowaniu w praktyce niesie za sobą dwie informacje: dobrą i złą.

- Dobra: sygnał o którym wiemy, że z pewnych naturalnych przyczyn ma widno zwarte może być całkowicie scharakteryzowany w formie cyfrowej.
- Zła: sygnał który ma niezwarłe widmo, w trakcie próbkowania ze skończoną dokładnością może produkować zaburzenia: w praktyce część widma znajdująca się poza zakresem próbkowania zostaje "wciśnięta" do tego zakresu. Zjawisko to nosi nazwę *aliasingu*.



Rysunek: Aliasing. Sygnał spróbkowany ze zbyt małą częstotliwością produkuje artefakt w postaci fali o niskiej częstotliwości, której w rzeczywistości w sygnale nie ma.



Jak przeciwdziałać zjawisku aliasingu?

- Zwiększyć częstotliwość próbkowania: czasem drogie a czasem niemożliwe...
- Obciąć wysokie częstotliwości na sygnale analogowym przed próbkowaniem
- Aplikować filtr dolnoprzepustowy na cyfrowym sygnale i zmniejszyć rozdzielczość: często jedyne wyjście gdy nie mamy dostępu do oryginalnego sygnału.



Powróćmy na chwilę do filtrów. Jak widzieliśmy filtry liniowe oraz inwariantne czasowo są reprezentowane splotem

$$c(k) = \sum_{i=0}^k a(i)b(k-i) \quad (11)$$

W wersji ciągłej splot wyraża się następująco:

$$c(x) = \int_a^b a(\tau)b(x-\tau)d\tau \quad (12)$$



Filtry splotowe są bardzo ważne, ponieważ:

- Umiemy je efektywnie implementować za pomocą FFT lub bezpośrednio gdy operator splotu jest bardzo zwarty dziedzinie czasu (wtedy wyliczenie splotu bezpośrednio może być efektywniejsze niż liczenie FFT)
- Cała masa użytecznych filtrów: dolno, górno-przepustowych, pasmowych wyraża się właśnie za pomocą splotu!



Sygnał dwuwymiarowy - obraz

No i dotarliśmy do obrazów, które możemy także traktować jako specyficzne sygnały:

- Obraz nieruchomy jest specyficznym sygnałem pozbawionym czasu (niezmiennym), który ma jednak cechy przestrzenne. W tym przypadku interesują nas własności przestrzenne, wszelkie filtry będą brały pod uwagę właśnie te cechy.
- Obraz ruchomy (film) możemy traktować jako
 - zbiór (uporządkowany) nieruchomych sygnałów przestrzennych
 - zbiór (uporządkowanych przestrzennie) sygnałów czasowych
 - sygnał trójwymiarowy przestrzennie, pozbawiony czasu

Zależnie od interpretacji możemy aplikować filtry w różnych kierunkach.



Podobnie jak w jednym wymiarze, możemy rozpatrywać transformatę Fouriera w większej ilości wymiarów:

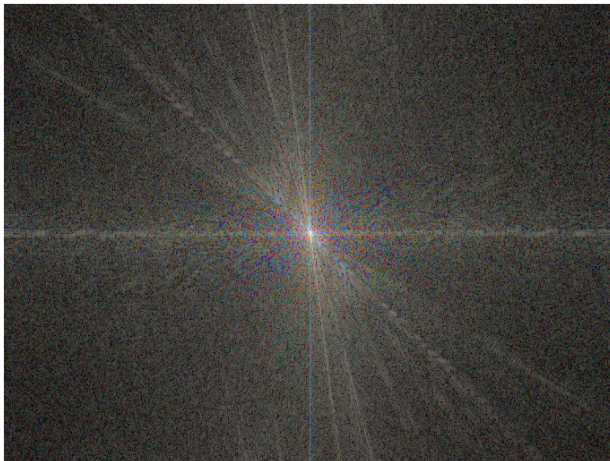
$$F(\omega_0, \omega_1) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t_0, t_2) e^{-2\pi i(t_0 \cdot \omega_0 + t_1 \cdot \omega_1)} d\omega_1 d\omega_0$$

W praktyce wzór ten jest łatwy - oznacza to, że tranformatę Fouriera dwuwymiarowego obrazu robimy następująco:

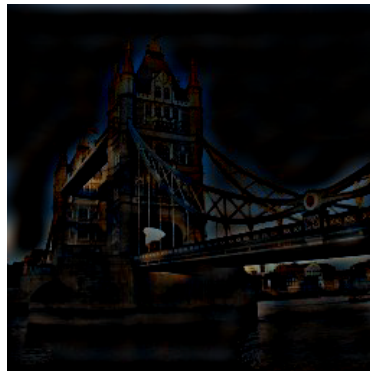
- Transformujemy po kolei wiersze (linie poziome) jak sygnał jednowymiarowy
- Następnie transformujemy powstałe kolumny, także jako sygnały jednowymiarowe



Rysunek: Obraz w dziedzinie przestrzennej.



Rysunek: Ten sam obraz w dziedzinie częstotliwości (widmo mocy).
Kompozycja trzech transformat osobno dla każdego koloru.



Rysunek: Informacje zawarte w niskich (lewo) częstotliwościach różnią się dalece od tych zawartych w wysokich (prawo). Wszelkie jednolite pola są reprezentowane w niskich częstotliwościach, szczegóły i krawędzie (bez wartości płaskich) znajdują reprezentacje w wysokich częstotliwościach.



W dwóch wymiarach odpowiednikiem filtra liniowego i inwariantnego czasowo jest ... filtr liniowy i inwariantny przestrzennie (jednorodny).

- Filtr taki jest wyznaczony jednoznacznie przez odpowiedź na impuls przestrzenny, który w tym przypadku jest deltą Diraca na płaszczyźnie (lub w przypadku dyskretnym pojedynczy zapalony pixel na czarnym tle).
- Odpowiedzią filtra na impuls jest pewien rozmazany obrazek. Z tego powodu odpowiedź impulsowa filtra dwuwymiarowego nazywa się czasem *funkcją rozmazania punktu* lub *PSF - Point Spread Function*.



Rysunek: Obraz rozmazany za pomocą eliptycznego operatora splotu.



Rozplatanie (deconvolution)

W pewnych zastosowaniach możemy natrafić na sprzyjające okoliczności:

- Otrzymujemy sygnał spleciony z pewnym operatorem
- Znamy PSF tego operatora
- Reprezentacja częstotliwościowa tego PSF jest odwracalna (nie ma zerowych częstotliwości w interesującym zakresie)

W takich okolicznościach możemy pokusić się o próbę "odwrócenia" splotu.



Rozplatanie (deconvolution)

Najkrócej rzecz ujmując schemat postępowania wygląda następująco:

- Transformujemy PSF do dziedziny częstotliwości
- Znajdujemy funkcję odwrotną ($\widehat{PSF}(\omega) * \widehat{PSF}^{-1}(\omega) = 1$) dla każdego ω w interesującym nas zakresie częstotliwości
- Aplikujemy $\widehat{PSF}^{-1}$ do bezpośrednio do transformaty obrazu, lub transformujemy $\widehat{PSF}^{-1}$ do dziedziny czasu (zależnie co jest bardziej efektywne i stabilne numerycznie)
- Otrzymujemy "odpleciony" obraz.



Rysunek: Obraz uzyskany w wyniku odwrócenia znanego operatora splotu poprzez filtr Wienera.



Rozplatanie - kubeł zimnej wody

Rozplatanie ma jednak ograniczone zastosowanie bo:

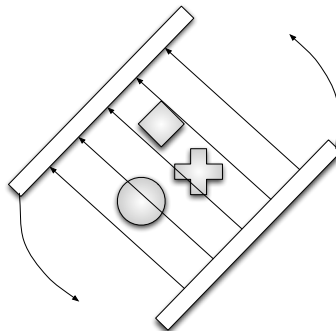
- Zazwyczaj nie znamy dokładnie PSF
- Szum na obrazie może nie być jednorodny przestrzennie (np. poruszone zdjęcie może mieć składową radialną)
- Szum może nie być liniowy (!) - aberracje optyczne często nie są liniowe!
- Rozplatanie jest bardzo podatne na niedokładności PSF i w związku z tym niestabilne numerycznie
- PSF może nie być odwracalne w dziedzinie częstotliwości!

... Co nie zmienia faktu, iż jest technika którą warto znać!

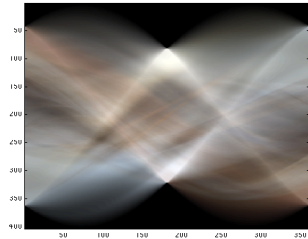
Transformata Radona

Istnieje jeszcze jedna ważna transformata dwuwymiarowego obrazu:

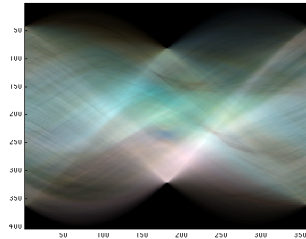
- Polega ona na "prześwietleniu" obrazu wzdłuż wszystkich kierunków
- W wyniku tej transformacji powstaje tak zwany "sinogram"
- Takie sinogramy produkuje np. tomograf



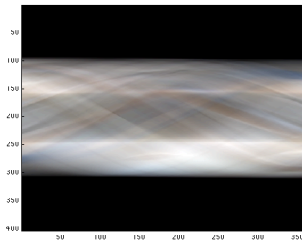
Rysunek: Schemat obrazowania tomograficznego.



Rysunek: Obraz oryginalny i jego sinogram (w RGB).



Rysunek: Obraz oryginalny i jego sinogram (w RGB).



Rysunek: Obraz oryginalny i jego sinogram (w RGB).



Czy transformata Radona jest odwracalna?

- Patrząc na pozwijany makaron sinogramu można odnieść wrażenie, iż tracimy sporo informacji
- Intuicja podpowiada na przykład, iż w ogólności nie da się odtworzyć funkcji na podstawie jej rzutów na osie
- Nawet mając wszystkie rzuty (kontury) nie da się odtworzyć funkcji która ma minimum lokalne!
- Tutaj jednak mamy nie rzuty a całki (prześwietlenia) i na wszystkie możliwe kierunki
- Wszystko to powoduje, że transformata Radona jest odwracalna! Łatwo to pokazać za pomocą transformaty Fouriera!



Projection-slice theorem

Twierdzenie (Projection-slice)

Transformata Fouriera projekcji (w sensie całkowym) funkcji na płaszczyźnie^a na pewną oś przechodzącą przez środek układu współrzędnych jest równa wycinkowi transformaty Fouriera całej funkcji wzdłuż linii przechodzącej przez środek układu współrzędnych w dziedzinie częstotliwości pod tym samym kątem co linia rzutowa we współrzędnych przestrzennych.

^aTwierdzenie ma odpowiedniki w wyższych wymiarach.

Mając sinogram, mamy transformaty Fouriera wzdłuż wycinków pod wszystkimi kątami, możemy zatem zrekonstruować całą transformatę Fouriera, a zatem i oryginalny obraz!



Projection-slice theorem - Dowód

DOWÓD. Bez straty ogólności rozumowania można założyć, że rzutujemy funkcję na oś x . Wtedy

$$p(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad (13)$$

Natomiast transformata Fouriera funkcji $f(x, y)$:

$$F(\omega_x, \omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-2i\pi(x\omega_x + y\omega_y)} dx dy \quad (14)$$



Projection-slice theorem - Dowód

DOWÓD. Wtedy wycinek transformaty to:

$$\begin{aligned} s(\omega_x) &= F(\omega_x, 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-2i\pi x \omega_x} dx dy = \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x, y) dy \right] e^{-2i\pi x \omega_x} dx = \\ &= \int_{-\infty}^{\infty} p(x) e^{-2i\pi x \omega_x} dx = \hat{p}(\omega_x) \end{aligned} \quad (15)$$

Co kończy dowód. ■



Owracanie Transformaty Radona - komentarz

- Wynikająca prosto z powyższego dowodu metoda odwracania transformaty Radona nie jest najlepsza, ze względu na znaczne niestabilności numeryczne.
- Zamiast tego w dwóch wymiarach² stosuje się metodę zwaną "filtered back-projection" ("filtrowana projekcja zwrotna" ?).

²Nie uogólnia się bezpośrednio do większej ilości wymiarów, chociaż istnieją pewne uogólnienia.



Filtrowana projekcja zwrotna

Przypomnijmy, że transformata Radona dana jest następująco:

$$g(s, \theta) = R(f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - s) dx dy \quad (16)$$

Zdefiniujmy operator wstecznej projekcji następująco:

$$B(g(s, \theta))[x, y] = \int_0^{\pi} g(x \cos \theta + y \sin \theta, \theta) d\theta \quad (17)$$

Operator ten "rozsmarowuje" każdy z plasterków transformaty Radona z powrotem na płaszczyznę a następnie sumuje je po wszystkich kątach. Operacja taka nie odwraca jednak transformaty, obraz będzie bardzo rozmażany.



Filtrowana projekcja zwrotna

Popatrzmy na złożenie:

$$\begin{aligned} B(g)[x, y] &= B(R(f))[x, y] = \\ &= \int_0^\pi \int_{-\infty}^\infty \int_{-\infty}^\infty f(u, v) \delta(u \cos \theta + v \sin \theta - x \cos \theta - y \sin \theta) du dv d\theta = \\ &= \int_{-\infty}^\infty \int_{-\infty}^\infty f(u, v) \left(\int_0^\pi \delta((u - x) \cos \theta + (v - y) \sin \theta) d\theta \right) du dv = \end{aligned} \quad (18)$$

Przypatrzmy się całce w nawiasie (dla przejrzystości zamieniono zmienne):

$$d(r, w) = \int_0^\pi \delta(r \cos \theta + w \sin \theta) d\theta \quad (19)$$

Zachowuje się ona miotła która zamiata całą płaszczyznę i zlicza punkty (tylko na przecięciu "miotły" z punktem δ jest niezerowa).
Tylko jak nadać temu "czemuś" wartość w punkcie?



Problem w tym, że funkcja d zdefiniowana poniżej nie jest funkcją w klasycznym sensie a raczej dystrybucją

$$d(r, w) = \int_0^\pi \delta(r \cos \theta + w \sin \theta) d\theta \quad (20)$$

Popatrzmy jak d zachowuje się wewnątrz całki przemnożone przez funkcję charakterystyczną I_A jakiegoś podzbioru płaszczyzny.

$$\int I_A(r, w) * d(r, w) dr dw = \int_A \int_0^\pi \delta(r \cos \theta + w \sin \theta) d\theta dr dw \quad (21)$$

Powyższa całka wyliczy pole zbioru A przemnożone przez "czas" który spędza na tym zbiorze promień wodzący! Jeśli będziemy próbować wartości tej całki nieskończenie małymi zbiorami o mierze ε , to wynik zawsze będzie odwrotnie proporcjonalny od odległości zbioru próbkowego od punktu $[0, 0]$. Dla zbioru A odległego od $[0, 0]$ o $l = \sqrt{l_r^2 + l_w^2}$ będzie to

$$\int I_A(r, w) * d(r, w) dr dw \approx \frac{\varepsilon}{l} = \frac{\varepsilon}{\sqrt{l_r^2 + l_w^2}} \quad (22)$$



Wracając do wzoru:

$$\begin{aligned} B(g)[x, y] &= B(R(f))[x, y] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u, v) \left(\int_0^{\pi} \delta((u-x) \cos \theta + (v-y) \sin \theta) d\theta \right) dudv = \end{aligned} \quad (23)$$

ponieważ jesteśmy pod całką, możemy dystrybucję zastąpić jej gęstością:

$$\begin{aligned} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u, v) \left(\int_0^{\pi} \delta((u-x) \cos \theta + (v-y) \sin \theta) d\theta \right) dudv = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u, v) \frac{1}{\sqrt{(u-x)^2 + (v-y)^2}} dudv = \\ &= f(x, y) ** \frac{1}{\sqrt{x^2 + y^2}} \end{aligned} \quad (24)$$

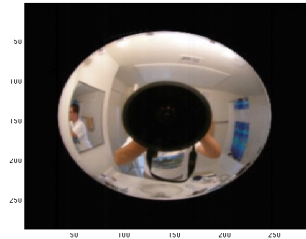
Mamy splot którego jądro znamy! Możemy więc wyostrzyć obraz znajdując odwrotność jądra splotu w dziedzinie częstotliwości.



Filtrowana projekcja zwrotna - jeszcze jeden trick

- Przypomnijmy sobie "Projection-slice theorem".
- Ponieważ dwuwymiarowe rozplatanie może być kosztowne, zauważmy iż możemy zmienić kolejność działań:
 - Najpierw zaaplikować stosowny filtr do plasterków transformaty Radona
 - Następnie wykonać projekcję wsteczną już bez potrzeby filtrowania

Algorytm ten jest prosty, stabilny i powoduje, że tomografy komputerowe działają w szpitalach od wielu lat. Ładne demo można zobaczyć tutaj: http://www.physics.ubc.ca/~mirg/home/tutorial/fbp_recon.html



Rysunek: Rekonstrukcje na podstawie sinogramów (filtrowana projekcja zwrotna).

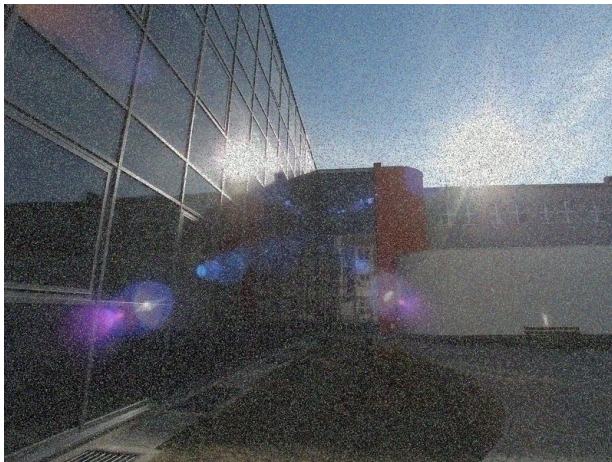


Eliminacja szumu punktowego

- W praktycznych zastosowaniach najczęstszym zakłóceniem obecnym w obrazach są przekłamane wartości pojedynczych próbek (pixeli)
- W pewnych przypadkach możemy mieć dodatkową informację o tym który pixel jest niewiarygodny (np. ustalone uszkodzenia rejestratora)
- W innych możemy być pozbawieni tej informacji (szum cieplny)
- Szum punktowy jest z pewnością cechą występującą w okolicach częstotliwości Nyquista, więc może po prostu zastosować filtr dolnoprzepustowy? Są lepsze rozwiązania.



Rysunek: Przykładowe zdjęcie.



Rysunek: To samo zdjęcie z zaaplikowanym 30% (!) szumem punktowym.

W przypadku gdy niewiarygodne pixele są znane, najlepiej (oczywiście) rozpiąć wielomian interpolacyjny na punktach otaczających i zinterpolować. Zważywszy, że przeważająca część energii przeciętnego obrazu skupiona jest w niskich częstotliwościach, ten trick prawie zawsze działa bardzo dobrze.

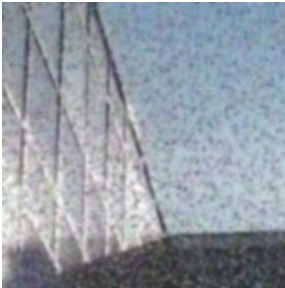


Rysunek: Fragment obrazu zaszumionego i zinterpolowanego w zaszumionych punktach.



Rysunek: Zinterpolowany obraz.

W przypadku gdy nie wiemy które pixele są błędne, wydawałoby iż nie pozostaje nic innego jak zastosować filtr rozmywający i wygubić trochę wysokich częstotliwości. Jak się okazuje zwykły filtr dolnoprzepustowy (uśredniający) daje tutaj słabe rezultaty, o wiele lepiej jest zastosować medianę!



Rysunek: Obraz uśredniony (lewy) i medianowany (prawy).



Dlaczego mediana działa lepiej?

- Uśrednienie zachowuje całkowitą "energię" obrazu, jedynie przesuwając jej masę w kierunku niskich częstotliwości
- Mediana natomiast potrafi wygubić energię (norma obrazu po zastosowaniu mediany może być istotnie mniejsza od oryginału), najlepiej jeśli filtr wytraca energię szumu!
- Zastanówmy się jak działa mediana na polu pixeli o tej samej wartości z jednym pixelem istotnie odbiegającym:
 - odstający pixel zostanie wygubiony (zastąpiony dominującą wartością)
 - pozostałe pixele zostaną nienaruszone!
 - norma (energia) obrazu się zmniejszyła!
- Uwaga! Mediana nie jest filtrem liniowym i NIE wyraża się w terminach splotu!



Aliasing w 2d

- Podobnie jak w przypadku sygnałów jednowymiarowych, tak i w 2d obowiązuje twierdzenie o próbkowaniu.
- Analogicznie pojawia się także problem aliasingu, gdy próbkowany obraz zawiera częstotliwości powyżej progu Nyquista.
- W 2d efekt ten może przyjmować ciekawą formę rozmaitych dziwnych fal przebiegających przez obraz - efekt ten nosi nazwę efektu Moire
- Problem ten rozwiązujemy tymi samymi metodami co w przypadku sygnału jednowymiarowego...



Rysunek: Obraz spróbkowany zbyt zgrubnie (ang. subsampled). Widać wyraźny efekt Moire.

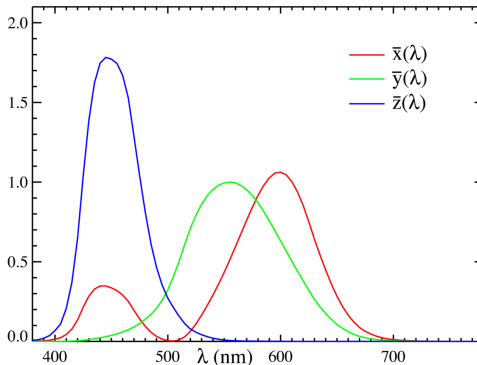


Rysunek: Ten sam obraz potraktowany filtrem dolnoprzepustowym przed próbkowaniem.



Kolory

- Jak wiadomo, ludzki aparat percepcji wzrokowej posiada receptory, których maksimum czułości przypada w trzech różnych częstotliwościach
- Są to długości fal: $\sim 445nm$, $\sim 545nm$ i $\sim 570nm$ którym odpowiadają odpowiednio wrażenia kolorystyczne niebieskiego, zielonego i czerwonego.
- Z tego względu obrazy przeznaczone do oglądania przez ludzi są próbkowane niezależnie w trzech pasmach, z tego względu każdy pixel składa się z trzech niezależnych intensywności.
- Uwaga: postrzegany kolor nie odpowiada składowi widmowemu światła - można na nieskończenie wiele sposobów konstruować światło które wywoła określone wrażenie kolorystyczne!
- Jednak światło o różnych widmach w różnym stopniu podlega dyfrakcji, więc nawet jeśli na pozór świetlówka i żarówka dają taki sam kolor światła, percepcja obiektów przez nie oświetlonych może być inna!



Rysunek: Stopień odpowiedzi ludzkiego aparatu percepcji na światło o określonej długości fali ustalony przez CIE (International Commission on Illumination) w 1931 roku. Zauważmy, że zielony i czerwony wcale nie są bardzo odległe!



Przestrzeń kolorów CIE XYZ

Dla ustalonych funkcji odpowiedzi możemy wyliczyć całkowitą odpowiedź na światło o widmie $I(\lambda)$:

$$\begin{aligned} X &= \int_0^{\infty} I(\lambda) \bar{x}(\lambda) d\lambda \\ Y &= \int_0^{\infty} I(\lambda) \bar{y}(\lambda) d\lambda \\ Z &= \int_0^{\infty} I(\lambda) \bar{z}(\lambda) d\lambda \end{aligned} \tag{25}$$

Zatem dla ustalonych funkcji odpowiedzi przestrzeń kolorów jest trójwymiarowa.



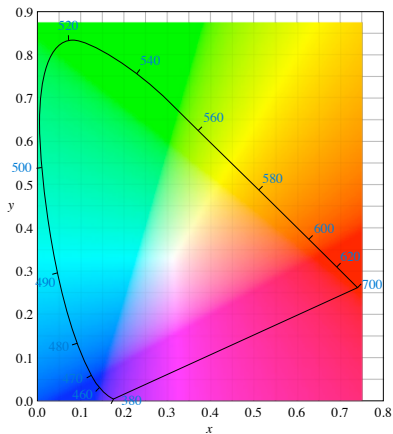
Całą przestrzeń kolorów można w sposób naturalny podzielić na składowe chromatyczności i natężenia:

$$\begin{aligned}x &= \frac{X}{X + Y + Z} \\y &= \frac{Y}{X + Y + Z} \\z &= \frac{Z}{X + Y + Z} = 1 - x - y\end{aligned}\tag{26}$$

Wtedy ustalając odpowiedź Y jako miarę całkowitego natężenia mamy reprezentację xyY :

$$\begin{aligned}X &= \frac{Y}{y}x \\Z &= \frac{Y}{y}(1 - x - y)\end{aligned}\tag{27}$$

Dla maksymalnej wartości Y dostajemy wtedy diagram chromatyczności w terminach x, y zwany diagramem CIE 1931

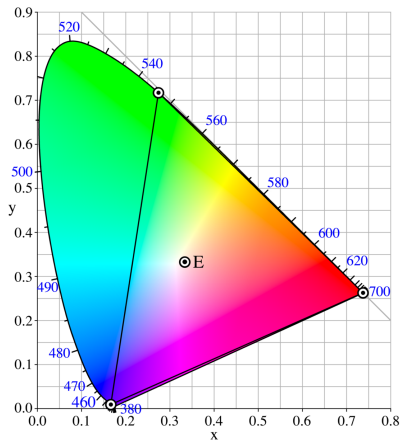


Rysunek: Diagram CIE 1931



Model CIE RGB

- Ze względów praktycznych (konstrukcji detektorów jak i urządzeń produkujących obraz) wygodnie jest rozpatrywać przestrzeń kolorów w której odpowiedź na sygnał maleje liniowo ze zmianą częstotliwości
- Takie modele rozpięte są na trzech kolorach (podstawowych), wszystkie inne kolory powstają jako suma ważona składowych podstawowych.
- W ramach ważenia należy także ustalić punkt bieli, czyli zestaw wag które produkują biel.
- Modele takie zazwyczaj rozpięte są na barwach podstawowych RGB (red, green, blue).
- Podzbiór kolorów danego modelu w przestrzeni CIE xyY nazywa się Gamutem.



Rysunek: Gamut CIE RGB



Modele RGB

- Dzisiejsze komputery zazwyczaj pracują w modelu *sRGB* o składowych $R = [0.64, 0.33]$, $G = [0.3, 0.6]$, $B = [0.15, 0.06]$ (w CIExyY) oraz punkcie bieli $D65 = [0.3128, 0.3290]$ odpowiadającemu temperaturze barwowej promieniowania ciała doskonale czarnego o temperaturze 6504K.
- Każde urządzenie takie jak monitor czy skaner ma swoje własne profile kolorystyczne i powinno być kalibrowane w celu uzyskania maksymalnej jakości obrazu. Niestety wiele osób nie ma o tym pojęcia...



Inne modele kolorów

Istnieje także wiele innych reprezentacji kolorów

- HSL - ze składowymi chromatycznymi Hue (odcień) i Saturation (natężenie) oraz składową luminacji Lightness
- HSV - podobny do HSL, jednak z inną reprezentacją luminacji
- CMYK - Cyan Magenta Yellow Black - model substraktywny używany w poligrafii
- YUV - ze składową luminacji luma (Y) oraz dwoma składowymi chromatycznymi UV - używany w telewizji (aby łagodnie przejść od telewizji czarno-białej do kolorowej) i kompresji JPEG
- CIE 1976 LAB ze składową luminacji L i składowymi chromatycznymi ab, opracowany ze szczególnym uwzględnieniem specyfiki ludzkiej percepcji



Zastosowania

- Po co w przetwarzaniu obrazu wszystkie te modele kolorów?
Nie lepiej ustalić jeden najlepszy i zapomnieć o innych?
- Rzecz w tym, że różne modele mają różne zalety i wady - nie ma najlepszego!
- Na koniec szybki przykład - przypomnijmy cechy transformaty Fouriera obrazu, a konkretnie to, że w wysokich częstotliwościach jest niewiele informacji kolorystycznej, jak to wykorzystać?



Zastosowania

- W przeciętnym zdjęciu widmo w wysokich częstotliwościach jest zazwyczaj praktycznie takie samo dla wszystkich składowych RGB.
- Wynika to z tego, iż w obrazach rzadko mamy krawędzie w jednej składowej kolorystycznej których nie ma w innych...
- Jeśli są jakieś różnice kolorystyczne w wysokich częstotliwościach, zazwyczaj ich źródłem jest szum cieplny
- Jeśli zaaplikujemy filtr wyostrzający (górnoprzepustowy) na składowych RGB mamy szansę wzmocnić szum!



Zastosowania

- Znacznie lepszym rozwiązaniem jest przejść do jednej z reprezentacji gdzie składowa luminancji jest oddzielona od składowych chromatycznych (najlepiej LAB lub HSL)
- Zaaplikować filtr górnoprzepustowy jedynie do składowej luminancji (na której szum cieplny jest uśredniony)
- Złożyć obraz z powrotem do RGB
- W ten sposób wzmocnimy wysokie częstotliwości, jednocześnie nie wzmacniając wariancji kolorystycznych tych częstotliwości!



Rysunek: Przykładowy obraz z lekkim szumem cieplnym.



Rysunek: Obraz potraktowany filtrem górnoprzepustowym w RGB.



Rysunek: Obraz potraktowany filtrem górnoprzepustowym składowej L przestrzeni LAB.



Spis treści
Wstęp
Sygnał jednowymiarowy
Sygnał dwuwymiarowy
Obraz kolorowy

Kolory i percepcja
Reprezentacje kolorów
Zastosowania



Dziekuję za uwagę!