

Transformace obrazu do frekvenční oblasti a jejich aplikace

Pavel Strachota, FJFI ČVUT v Praze

Abstract. Motivated by the algorithms of lossy image compression, this paper introduces the mathematics and applications of the discrete Fourier transform in a way accessible to high school students. Both the theoretical treatment and some visually appealing practical applications are presented.

Úvod

Většina z nás dnes rutinně používá digitální fotoaparát v mobilu a pořizuje s ním fotografie a videa, která sdílí přes internet. Z internetu nebo ze svého vlastního datového úložiště sledujeme filmy a posloucháme hudbu. Zvykli jsme si říkat, že si posíláme ty „JPEgy“, posloucháme „MP-trojky“ a sledujeme „MP-čtyřky“ (pravda, to poslední se zase tak často neříká). Abychom to mohli dělat, musíme mít na počítači nainstalované ty správné „kodeky“ (na mobilu to zpravidla až tolik neřešíme, prostě to funguje).

Zkratky JPEG [1], MP3 [2], MP4 [3] a mnoho dalších podobných odkazují na formáty ukládání (kódování) audiovizuálních dat tak sofistikovaným způsobem, aby výsledné datové soubory zabíraly co nejméně místa, a přitom bylo možné z nich rekonstruovat (dekódovat) původní obrazový nebo zvukový záznam. Kombinací slov kódování a dekodování pak vznikl výraz „kodek“, označující implementaci těchto algoritmů, a to typicky ve formě programové knihovny využitelné více různými aplikacemi.

Oproti objemu původních dat jsou výsledné soubory někdy i o několik řádů menší, čehož lze na rozdíl od univerzálních kompresních algoritmů (např. známého ZIPu) dosáhnout použitím ztrátové komprese. To znamená, že rekonstruovaná data nejsou (z hlediska uchované informace, tj. hodnot jednotlivých bajtů) identická s daty původními, ale subjektivně na pohled, resp. na poslech člověk pozná jen nepatrný rozdíl. Dnešní algoritmy ztrátové komprese šikovně využívají vlastností lidského vnímání a z původních dat zachovávají jen takovou informaci, která je pro smysly člověka podstatná. Klíčovým nástrojem k dosažení tohoto cíle jsou matematické transformace, které umožní takzvaně „převést data z

33 oblasti prostorové do frekvenční“. Například u transformovaného obrázku
34 již data neukazují, na kterém místě je která barva, ale z jak vysokých
35 vln o různých frekvencích se obrázek skládá.

36 Cílem první části našeho článku je vysvětlit a matematicky dokázat
37 princip těchto transformací a poté prakticky demonstrovat jejich funkci.
38 U čtenáře nepředpokládáme vysokoškolské znalosti matematiky, a proto
39 samotnému důkazu předchází shrnutí některých použitých pojmů a kon-
40 ceptů, které na první pohled s tématem přímo nesouvisí. V druhé části
41 se vrátíme k algoritmům ztrátové komprese obrazu a ukážeme si, ja-
42 kým způsobem se v nich využívají transformace do frekvenční oblasti.
43 Nechybí praktické příklady demonstrující efekt ztrátové komprese na re-
44 álné obrázky.

45 **Reprezentace rastrového obrazu**

46 Obrazová nebo zvuková data lze matematicky vyjádřit jako funkci
47 závislou na čase a prostorových souřadnicích. Například rovinný šedotó-
48 nový (tj. „černobílý“) obdélníkový obraz je dán funkcí f , která pro každý
49 bod (x, y) z tohoto obdélníku dodá číslo vyjadřující intenzitu (světlost)
50 $f(x, y)$ v tomto bodě v rozsahu od 0 (černá) do 1 (bílá). V případě
51 rastrového obrazu složeného z obrazových bodů (pixelů) je však situace
52 jednodušší. Předpokládáme-li, že pixely mají čtvercový tvar a obrázek
53 má velikost $N \times M$ pixelů, je funkce f po částech konstantní (s kon-
54 stantní hodnotou na ploše každého pixelu) a může teoreticky¹⁾ nabývat
55 nejvýše MN hodnot. Veškerou informaci o obrázku lze proto uložit do
56 matice (tabulky čísel)

$$\mathbf{f} = \begin{pmatrix} f_{0,0} & f_{0,1} & \cdots & f_{0,N-1} \\ f_{1,0} & f_{1,1} & \cdots & f_{1,N-1} \\ \vdots & \vdots & & \vdots \\ f_{M-1,0} & f_{M-1,1} & \cdots & f_{M-1,N-1} \end{pmatrix}, \quad (1)$$

57 kde $f_{j,k}$ vyjadřuje intenzitu pixelu, který je v obraze j -tý shora a k -
58 tý zleva, přičemž začínáme indexovat od nuly. Takto definovaná data
59 lze transformovat do frekvenční oblasti pomocí dvourozměrné *diskrétní*

¹⁾Víme, že na počítači je i počet různých barev omezen. Např. v režimu True Color lze vybírat z celkem cca 16,7 milionu různých barev, ale pouze 256 z nich jsou odstíny šedi.

60 *Fourierovy transformace* (DFT)²⁾. Než si ukážeme jak, připomeneme si
61 pár věcí z matematiky a trochu se procvičíme.

62 Součet N členů geometrické posloupnosti

63 Uvažujme posloupnost N čísel $a_0, a_1, a_2, \dots, a_{N-1}$. Jejich součet zna-
64 číme

$$\sum_{n=0}^{N-1} a_n = a_0 + a_1 + a_2 \cdots + a_{N-1},$$

65 znak Σ (velké řecké písmeno sigma) se nazývá *suma* a písmeno n je
66 *sčítací index*, který při sčítání dané N -tice čísel nabývá postupně hodnot
67 $0, 1, 2, \dots, N - 1$.

Jednou z důležitých číselných posloupností je geometrická posloup-
nost, pro níž $a_n = q^n$ a q je dané reálné číslo, tzv. kvocient. Součet s_N
prvních N členů geometrické posloupnosti najdeme snadno, pokud se
nejprve zeptáme, kolik je $(1 - q) s_N$. Dostaneme totiž

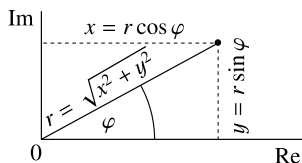
$$\begin{aligned} (1 - q) s_N &= (1 - q) \sum_{n=0}^{N-1} q^n = \sum_{n=0}^{N-1} q^n - q \sum_{n=0}^{N-1} q^n = \sum_{n=0}^{N-1} q^n - \sum_{n=0}^{N-1} q^{n+1} \\ &= \sum_{n=0}^{N-1} q^n - \sum_{j=1}^N q^j = 1 - q^N. \end{aligned} \quad (2)$$

68 Úprava na druhém řádku v (2) spočívá v tzv. posunutí indexu, kdy
69 definujeme $j = n + 1$, takže jestliže se index n pohyboval od 0 do $N - 1$,
70 tak j půjde od 1 do N . V tu chvíli ale vidíme, že v obou sumách sčítáme
71 pro $n = j = 1, 2, \dots, N - 1$ stejná čísla, která se navzájem odečtou.
72 První suma má ovšem navíc ještě sčítanec $q^0 = 1$ (pro $n = 0$) a druhá
73 má zase navíc q^N (pro $j = N$). Z této úvahy již plyne poslední rovnost
74 v (2). Vyjádřením s_N z (2) pak pro $q \neq 1$ okamžitě dostáváme hledanou
75 hodnotu

$$s_N = \frac{1 - q^N}{1 - q}, \quad (3)$$

76 čehož později využijeme.

²⁾ neboli Fourierovy transformace aplikované na *diskrétní data*, tj. hodnoty defino-
vané na konečné množině oddělených bodů. Mimo to existuje i (spojitá) Fourierova
transformace, která transformuje funkce definované na množině všech reálných čísel.



Obr. 1: Komplexní číslo $z = x + iy$ jako bod v komplexní rovině a jeho polární souřadnice.

77 Komplexní čísla

78 Rozumíme již tedy trochu sumám, ovšem stále jsme se nedostali k
 79 věci. Ještě jednou (ale slibuji, že už naposledy) se zdržíme, tentokrát u
 80 komplexních čísel. Popisovat užitečnost komplexních čísel je mimo rámec
 81 tohoto textu, ale při troše zjednodušení lze říci, že v oboru komplexních
 82 čísel mají řešení všechny kvadratické rovnice, včetně těch se záporným
 83 diskriminantem. Komplexní čísla jsou totiž ve tvaru

$$z = x + iy, \text{ kde } x, y \text{ jsou reálná čísla a } i^2 = -1.$$

84 Číslo i říkáme *imaginární jednotka*, x je reálná část a y je imaginární
 85 část komplexního čísla z . Číslo $\bar{z} = x - iy$ nazýváme číslo komplexně
 86 sdružené k číslu z .

Každé komplexní z si lze také představit jako bod v rovině (tzv. *komplexní rovině*) o souřadnicích (x, y) , jak je vidět na obr. 1. Polohu tohoto bodu lze určit i v polárních souřadnicích, tj. jeho vzdáleností od počátku r a úhlem φ , který svírá jeho průvodič (spojnice s počátkem souřadné soustavy) s osou x (tzv. reálnou osou, protože čísla, pro která $y = 0$, jsou reálná). Zřejmě platí

$$\begin{aligned} x &= r \cos \varphi, \\ y &= r \sin \varphi, \\ r &= \sqrt{x^2 + y^2}. \end{aligned}$$

87 Číslu r říkáme *absolutní hodnota* komplexního čísla z (značíme $|z|$) a
 88 číslu φ *argument* komplexního čísla z .

89 Ukazuje se, že definice většiny elementárních funkcí reálné proměnné
 90 (např. $\sin x$, $\cos x$, e^x) lze vhodně rozšířit i do komplexního oboru. Pro tři
 91 právě zmíněné funkce a libovolné komplexní číslo z platí velmi důležitý

92 vztah

$$e^{iz} = \cos z + i \sin z, \quad (4)$$

93 kterému se říká Eulerův vzorec. Díky němu můžeme psát

$$z = x + iy = r (\cos \varphi + i \sin \varphi) = re^{i\varphi},$$

94 a výrazu na pravé straně říkáme *goniometrický tvar* komplexního čísla z .

95 Z něj je například vidět, co se děje, když násobíme komplexní čísla $z_1 =$

96 $r_1 e^{i\varphi_1}$ a $z_2 = r_2 e^{i\varphi_2}$. Dostáváme

$$z_1 z_2 = r_1 e^{i\varphi_1} r_2 e^{i\varphi_2} = r_1 r_2 e^{i(\varphi_1 + \varphi_2)},$$

97 takže součin má absolutní hodnotu rovnou součinu absolutních hodnot

98 činitelů a argument rovný součtu argumentů jednotlivých činitelů³⁾.

99 Diskrétní Fourierova transformace

Konečně máme vše připraveno pro definici DFT. Začneme pouze v jednorozměrném případě, kdy zdrojová data představují posloupnost N čísel f_0, f_1, \dots, f_{N-1} . Výsledkem DFT je jiná posloupnost N čísel F_0, F_1, \dots, F_{N-1} , tzv. *Fourierův obraz* nebo *spektrum* původní posloupnosti, jejíž n -tý prvek je definován vztahem

$$F_n = \sum_{k=0}^{N-1} f_k e^{-2\pi i k n / N} \quad \text{pro } n = 0, 1, \dots, N-1. \quad (5)$$

Čísla f_k mohou být reálná i komplexní, ale čísla F_n jsou obecně komplexní. Velmi důležité je, že ze znalosti Fourierova obrazu lze vždy získat původní posloupnost, a to pomocí inverzní DFT (iDFT), která má tvar

$$f_k = \frac{1}{N} \sum_{n=0}^{N-1} F_n e^{2\pi i k n / N} \quad \text{pro } k = 0, 1, \dots, N-1. \quad (6)$$

100 Toto tvrzení si snadno dokážeme. Zvolíme-li pevné k a dosadíme-li do

101 pravé strany (6) tvar F_n ze vztahu (5), dostaneme

$$\frac{1}{N} \sum_{n=0}^{N-1} F_n e^{2\pi i k n / N} = \frac{1}{N} \sum_{n=0}^{N-1} \left(\sum_{j=0}^{N-1} f_j e^{-2\pi i j n / N} \right) e^{2\pi i k n / N}, \quad (7)$$

3) ... z čehož pak plyne tzv. Moivreova věta

$$(\cos x + i \sin x)^n = \cos(nx) + i \sin(nx).$$

102 kde jsme museli zvolit jiné označení sčítacího indexu uvnitř velké závorky
 103 (j místo k), protože k už je pevně dané číslo. Obecně celkem N^2 sčítanců
 104 obsažených v tomto součtu lze sčítat v libovolném pořadí, což využijeme
 105 k úpravě (7) do tvaru

$$\sum_{j=0}^{N-1} f_j \frac{1}{N} \sum_{n=0}^{N-1} e^{-2\pi i j n / N} e^{2\pi i k n / N} = \sum_{j=0}^{N-1} f_j \frac{1}{N} \underbrace{\sum_{n=0}^{N-1} e^{2\pi i (k-j) n / N}}_{\text{ozn. } s_j}, \quad (8)$$

106 kde první (vnější) suma sčítá přes sčítací index j a vnitřní suma, kterou
 107 jsme označili jako s_j , sčítá přes n (v (7) to totiž bylo opačně). Navíc
 108 f_j bylo možné z vnitřní sumy vytknout, protože je to číslo nezávislé
 109 na n . Pro $j = k$ vnitřní suma obsahuje N jedniček ($e^0 = 1$), takže platí
 110 $s_k = N$. Pro $j \neq k$ si uvědomíme, že s_j je pouze součtem prvních N
 111 členů geometrické posloupnosti s kvocientem

$$q = e^{2\pi i (k-j) / N}.$$

112 Dle (3) tedy platí

$$s_j = \sum_{n=0}^{N-1} e^{2\pi i (k-j) n / N} = \frac{1 - q^N}{1 - q} = \frac{1 - e^{2\pi i (k-j)}}{1 - e^{2\pi i (k-j) / N}} = 0.$$

113 Poslední rovnost je vidět z Eulerova vzorce (4), neboť $(k - j)$ je celé
 114 číslo, a funkce \sin , \cos jsou periodické s periodou 2π , takže

$$e^{2\pi i (k-j)} = \cos(2\pi (k - j)) + i \sin(2\pi (k - j)) = 1 + i \cdot 0.$$

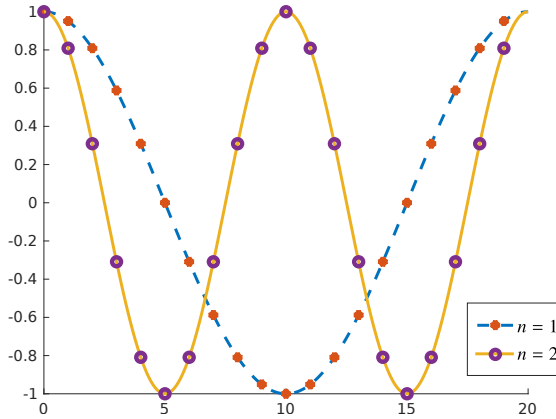
115 Tím jsme zjistili, že hodnota (8), což je hodnota pravé strany rovnosti
 116 (6), je skutečně rovna f_k .

117 Abychom pochopili význam čísel F_n , vyjádříme si je v goniometrickém
 118 tvaru jako

$$F_n = r_n e^{i\varphi_n}$$

a dosadíme do definičního vztahu iDFT (6). Dostaneme

$$\begin{aligned} f_k &= \frac{1}{N} \sum_{n=0}^{N-1} r_n e^{i\varphi_n} e^{2\pi i k n / N} = \sum_{n=0}^{N-1} \frac{r_n}{N} e^{i(2\pi k n / N + \varphi_n)} \\ &= \sum_{n=0}^{N-1} \frac{r_n}{N} (\cos(2\pi k n / N + \varphi_n) + i \sin(2\pi k n / N + \varphi_n)). \end{aligned} \quad (9)$$



Obr. 2: Grafy funkce $\cos(2\pi xn/N)$ pro $N = 20$ a $n = 1, 2$ a její vzorky v N bodech, tj. hodnoty v bodech $x = k = 0, 1, \dots, N - 1$, které (když $\varphi_n = 0$) vystupují ve vztahu (9).

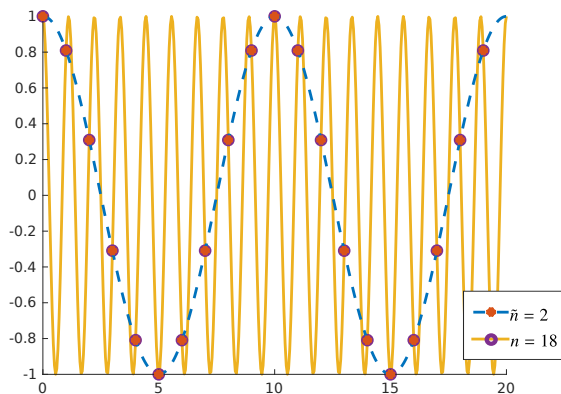
119 Z toho je vidět, že absolutní hodnoty čísel F_n (tj. čísla r_n) jsou až na
 120 faktor $1/N$ rovny amplitudám kosinových, resp. sinových vln, které jsou
 121 vzorkovány v bodech $(2\pi kn/N)$ pro $k = 0, 1, \dots, N - 1$. Argumenty čísel
 122 F_n (tj. čísla φ_n) pak definují jejich fázové posunutí. Sečtením všech tak-
 123 kových vln vzniká původní posloupnost f_k . Přitom pro $n = 1$ projdeme
 124 pro $k = 0, 1, \dots, N - 1$ právě přes jednu periodu funkcí \cos , resp. \sin , pro
 125 $n = 2$ projdeme přes 2 periody atd. (viz obr. 2). Index n se proto nazývá
 126 *frekvence*.

127 Je docela zajímavé pozorovat, co se děje pro vysoké frekvence, tj. když
 128 n je blízko počtu členů posloupnosti N . Definujme si $\tilde{n} = N - n$, kde
 129 $n > N/2$. Snadný výpočet ukáže, že platí

$$e^{2\pi i k n / N} = e^{2\pi i k (N - \tilde{n}) / N} = \underbrace{e^{2\pi i k N / N}}_{=1} e^{-2\pi i k \tilde{n} / N} = e^{-2\pi i k \tilde{n} / N} = \overline{e^{2\pi i k \tilde{n} / N}}. \quad (10)$$

130 V posledním kroku jsme využili Eulerův vzorec a definici komplexně
 131 sdruženého čísla. To znamená, že pro „velké“ $n > N/2$ má reálná část
 132 výrazu $e^{2\pi i k n / N}$ (tj. $\cos(2\pi i k n / N)$) stejnou hodnotu jako pro „malé“
 133 $\tilde{n} = N - n < N/2$, jak je vidět na obr. 3. Imaginární části pak mají
 134 navzájem opačná znaménka.

135 Jestliže jsou navíc f_k reálná čísla (což při využití DFT pro transfor-



Obr. 3: Graf funkcí $\cos(2\pi xn/N)$ a $\cos(2\pi x\tilde{n}/N)$ pro $N = 20$, $n = 18$ a $\tilde{n} = N - n = 2$. I když jsou výsledné funkce evidentně odlišné, jejich vzorky v bodech $x = k = 0, 1, \dots, N - 1$ jsou stejné.

136 mace obrazu bude splněno), plyne z (10) a definičního vztahu (5) také
137 vztah

$$F_n = \overline{F_{N-n}}. \quad (11)$$

138 Po aplikaci DFT na N reálných hodnot f_0, \dots, f_{N-1} tedy dostáváme
139 (pro N sudé) jen $N/2$ nezávislých hodnot $F_0, \dots, F_{N/2-1}$ a zbytek lze
140 dopočítat z (11). To není překvapující, protože F_n jsou komplexní a
141 „vydají“ za dvě reálná čísla, takže v nich je dost informací na to, aby
142 bylo možné původní N -tici rekonstruovat pomocí iDFT dle vztahu (6).

143 Spektrum digitálního obrazu a jeho úpravy

Nyní si ukážeme zajímavé aplikace DFT na obrazová data. Budeme potřebovat dvourozměrnou DFT, která vznikne dvěma aplikacemi jedno-rozměrné DFT na matici \mathbf{f} definovanou dle (1). Nejprve transformujeme každý j -tý řádek obrazu dle

$$\tilde{F}_{j,n} = \sum_{k=0}^{N-1} f_{j,k} e^{-2\pi i kn/N}.$$

Poté aplikujeme DFT na výsledné hodnoty $\tilde{F}_{j,n}$ po sloupcích:

$$\begin{aligned}
 F_{m,n} &= \sum_{j=0}^{M-1} \tilde{F}_{j,n} e^{-2\pi i j m / M} = \sum_{j=0}^{M-1} \left(\sum_{k=0}^{N-1} f_{j,k} e^{-2\pi i k n / N} \right) e^{-2\pi i j m / M} \\
 &= \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} f_{j,k} e^{-2\pi i (k n / N + j m / M)}. \tag{12}
 \end{aligned}$$

144 Kdybychom postupovali nejprve po sloupcích a poté po řádcích, získali
 145 bychom totéž.

146 Zvolme si obrázek, na němž jsou vidět výrazné pravidelné pruhy, na-
 147 příklad ten na obr. 4a. Po aplikaci DFT si zobrazíme absolutní hodnoty
 148 spektra též jako obrázek, kde světlost jednotlivých bodů je úměrná hod-
 149 notám $|F_{m,n}|$ ⁴⁾. Kvůli symetrii (11), která nyní platí v obou rozměrech,
 150 je však šikovně přesunout pravou polovinu spektra doleva a dolní polo-
 151 vinu nahoru, takže dostaneme výsledek jako na obr. 4b, kde zastoupení
 152 nejnižších frekvencí ukazují body uprostřed. Hned si všimneme světlej-
 153 šších svislých pruhů. Ty dva nejbližše středu odpovídají frekvenci střídání
 154 pruhů na stěně. Pruhy jsou ale ostře ohraničené, tj. jejich intenzita se
 155 nemění plynule jako hodnoty funkce sinus nebo kosinus. K získání ost-
 156 rého barevného přechodu jsou totiž zapotřebí i vyšší frekvence, kterým
 157 odpovídají pruhy dále od středu. Abychom si tuto myšlenku ověřili, od-
 158 stranili jsme všechny zmíněné pruhy ze spektra (příslušné hodnoty jsme
 159 nastavili na nulu), což je vidět na obr. 4d. Nakonec jsme provedli iDFT
 160 a výsledkem je obr. 4c⁵⁾. Pruhy skutečně z velké části zmizely, ale naše
 161 úpravy způsobily jako vedlejší efekt i částečné poškození zbytku obrázku.

162

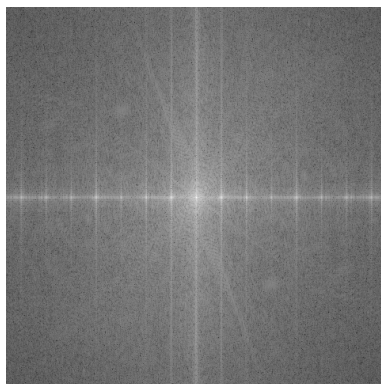
163 Pro demonstraci schopností DFT si ještě vyrobíme novou variantu
 164 optického klamu „Marilyn nebo Einstein“, která byla před časem velmi
 165 populární na internetu [4]. Jako zdrojová data nám poslouží fotografie
 166 obličejů Marilyn Monroe (obr. 5a) a Alberta Einsteina (5b), avšak zá-
 167 měrně jiné, než byly použity v původní verzi iluze. Po získání spektra

⁴⁾Přesněji řečeno, světlost je na obr. 4b úměrná logaritmu $|F_{m,n}|$, abychom viděli jako dostatečně světlé i amplitudy vyšších frekvencí. Se zvyšováním frekvence totiž u každého „rozumného“ obrázku její amplituda prudce klesá, a to i o několik řádů.

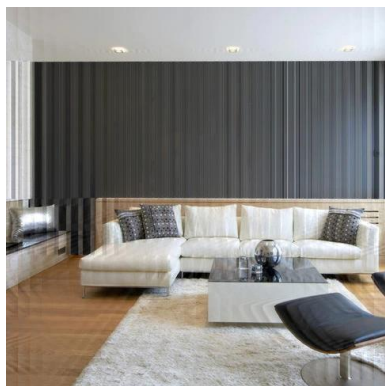
⁵⁾Pozornému čtenáři neujde, že obrázek je barevný, přičemž jsme se dosud bavili jen o šedotónových obrázcích. Barva ale není pro nás až taková věda. Stejně úpravy spektra jsme totiž provedli pro všechny tři barevné kanály R,G,B a výsledek pak opět složili dohromady.



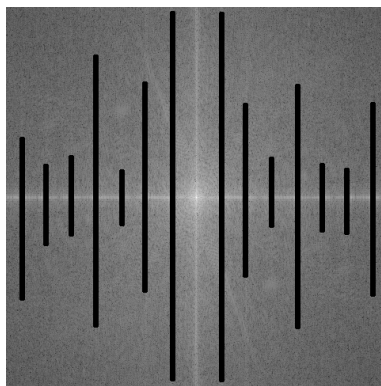
(a) Původní obrázek



(b) Spektrum původního obrázku



(c) Rekonstruovaný obrázek



(d) Modifikované spektrum

Obr. 4: Odstranění některých frekvencí ve frekvenční oblasti a rekonstrukce obrázku z modifikovaného spektra.

168 fotografie Marilyn v něm vynulujeme amplitudy všech vyšších frekvencí
169 mimo malý kroužek okolo středu (okolo nulové frekvence), neboli apliku-
170 jeme *nízkofrekvenční filtr*. Po provedení iDFT získáme obr. 5c. Je vidět,
171 že z fotografie zmizely všechny detaily a je velmi „rozmazaná“. V pravém
172 horním rohu je pro představu zobrazena i aplikace nízkofrekvenčního fil-
173 tru na spektrum, ovšem pro ilustraci je místo úplného vynulování použito
174 jen ztmavení. Spektrum fotografie Einsteina zpracujeme zcela opačně.
175 Aplikujeme *vysokofrekvenční filtr*, neboli vynulujeme jen malé kruhové
176 okolí středu. Výsledek je na obr. 5d, kde naopak zůstaly kontury a tvary,
177 ale zmizely barvy. Výslednou iluzi na obr. 6 pak získáme sečtením obou
178 obrázků 5c a 5d. Na její princip přijdete snadno: pozorujeme-li obrázek
179 z velké dálky (případně si sundáme brýle, případně si obrázek v PDF na
180 počítači zmenšíme), detailní kontury nevidíme a zdá se nám, že na ob-
181 rázku je Marilyn. Pokud se však podíváme dostatečně zblízka, zaujmu
182 nás detailní kontury a na obrázku rozpoznáme Alberta Einsteina s vy-
183 plazeným jazykem.

184 Ztrátová komprese obrazu

185 Vraťme se nyní k algoritmům ztrátové komprese obrazu, kterými jsme
186 naši diskusi začali. Vzhledem k tomu, o jak komplikovanou a rozsáhlou
187 problematiku se dnes jedná, se omezíme na stručný a ne zcela přesný
188 popis.

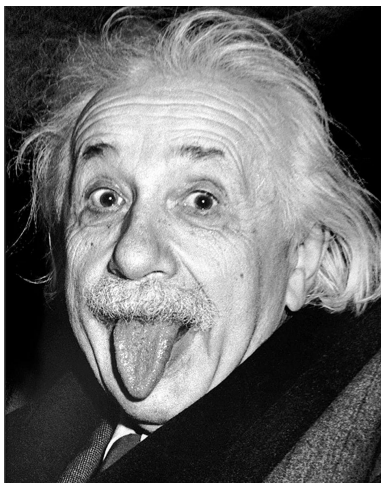
189 Komprese do více než 30 let starého a přitom stále populárního for-
190 mátu JPEG začíná transformací barevných kanálů R, G, B, jejímž vý-
191 sledkem je pro každý pixel informace o jeho světlosti (intenzitě) a o jeho
192 odstínu. Vzhledem k tomu, že lidské oko mnohem snáze rozezná rozdíly v
193 intenzitě než rozdíly v barevném odstínu, neukládají se hodnoty odstínu
194 pro každý pixel samostatně, ale jen pro bloky o velikosti 2×2 pixely.

195 Poté nastupuje převod dat o intenzitách i o odstínech barev do frek-
196 venční oblasti, a to samostatně a vždy po čtvercích velikosti 8×8 . Nepo-
197 užívá se přímo DFT, nýbrž tzv. diskrétní kosinová transformace (DCT)
198 [5], jejíž tvar lze odvodit aplikací DFT na data rozšířená pomocí vhodné
199 symetrie (pro zájemce jsou technické detaily shrnuty v dodatku). Ve vý-
200 sledku DCT pracuje jen s funkcí kosinus a výhodou je, že po transformaci
201 reálných dat získáváme opět reálné koeficienty. Vlastnosti výsledku jsou
202 však stejné, jako jsme si demonstrovali v předchozích experimentech s
203 DFT: amplitudy kosinových vln velmi rychle klesají s frekvencí.

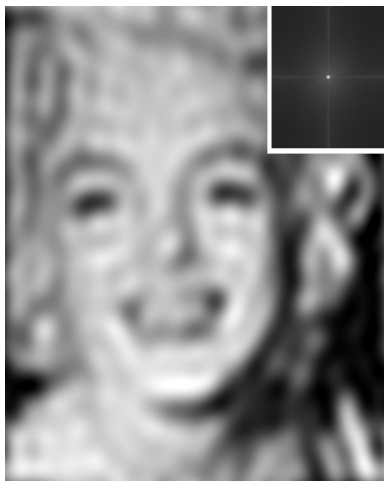
204 Ztrátovost komprese spočívá ve vhodném zaokrouhlení těchto ampli-
205 tud, přičemž se spoléháme na to, že více či méně z nich se zaokrouhlí



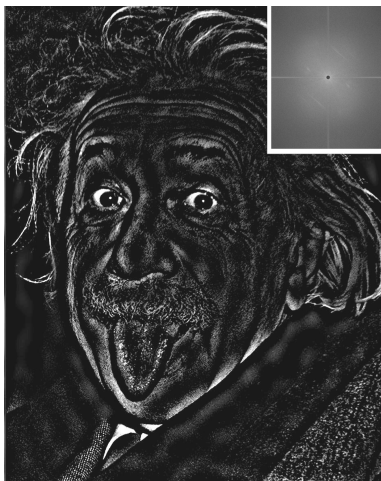
(a) Původní fotografie Marilyn Monroe



(b) Původní fotografie Alberta Einsteina



(c) Marilyn po aplikaci nízkofrekvenčního filtru



(d) Einstein po aplikaci vysokofrekvenčního filtru

Obr. 5: Příprava optického klamu „Marilyn nebo Einstein?“.



Obr. 6: Výsledná podoba optického klamu „Marilyn nebo Einstein?“.

206 na nulu. Různou úroveň zaokrouhlování může uživatel řídit pomocí ko-
207 eficientu kvality, který je vyjádřen v procentech. Z každého bloku 8×8
208 tak zbude jen několik málo nenulových hodnot, které se vhodně seřadí
209 za sebe a podstoupí bezztrátovou kompresi pomocí dalších algoritmů, o
210 nichž tento článek nepojednává. Při dekódování pomocí inverzní DCT se
211 pak z malého počtu nenulových hodnot obnoví celý čtverec 8×8 , jehož
212 hodnoty se pro „rozumné“ obrázky příliš neliší od originálu.

213 Výsledná úroveň komprese závisí na nastavení kvality i na složitosti
214 obrázku, což si autor zkusil na svých nedávných fotografiích. Při rozlišení
215 5792×4344 pixelů (cca 25 megapixelů) a ukládání barevných hodnot R,
216 G, B ve třech bytech pro každý pixel by každý takový obrázek zabíral
217 v paměti cca 75 MB. Fotografie venkovské krajiny se strništěm (obr.
218 7a), kde je k rozeznání každé stéblo, má ve formátu JPEG s nastavením
219 kvality na 80% cca 5 MB, což je 1/15 původního objemu. Ovšem snímek
220 městské zástavby (obr. 7b) má jen necelé 3 MB, což je dokonce 1/25
221 původního objemu dat.

222 V případě fotografií je obvykle nastavení kvality na 80% dostatečné,
223 neboť neumožňuje pouhým okem rozeznat jakýkoli rozdíl oproti nekom-
224 primovanému obrazu. Abychom v tomto článku dokázali demonstrovat
225 viditelné snížení kvality, uložili jsme výřez ze zmíněné fotografie krajiny
226 s nastavením kvality na pouhých 5%. Výsledek na obr. 8 ukazuje⁶⁾, že v
227 mnoha případech (zejména na obloze) jsou bloky 8×8 vykresleny kon-
228 stantní barvou. To znamená, že po zaokrouhlení koeficientů DCT zůstala
229 v každém z těchto bloků jediná nenulová hodnota. Zároveň je značně re-
230 dukován i počet barev.

231 Ve skutečnosti jsou však rozdíly oproti originálu přítomné pro jaké-
232 koliv nastavení kvality, což jsme se pokusili názorně ověřit pomocí volně
233 dostupného grafického editoru GIMP⁷⁾. Uložili jsme fotografii na obr.
234 7a ve formátu JPEG s různým nastavením kvality a výsledek pak ode-
235 četli od originálu, což je rovněž JPEG s kvalitou 98%. Aby byly rozdíly
236 lépe viditelné, zvýraznili jsme je úpravou jasu a kontrastu (ve všech
237 případech stejnou, aby bylo možné rozdíly porovnávat). Několik takto
238 zpracovaných „rozdílových obrazů“ je vidět na obr. 9, který ukazuje,
239 že se zvyšováním kvality dochází nejen ke zmenšování rozdílů, ale také
240 k zjemňování struktury artefaktů ztrátové komprese. Průměrná inten-

⁶⁾Zobrazuje-li si čtenář tento článek v prohlížeči PDF souborů na obrazovce počítače, tabletu či mobilu, doporučujeme zvětšit si obr. 8 (a stejně tak i demonstrace na obr. 9, 10 a 11) a pozorně si prohlédnout detaily.

⁷⁾<https://www.gimp.org/>



(a) kvalita 80%, velikost 4900 kB



(b) kvalita 80%, velikost 2840 kB

Obr. 7: Rozdílný kompresní poměr formátu JPEG v závislosti na obsahu snímků s rozlišením 5792×4344 pixelů.

241 zita rozdílových obrazů společně s velikostí výsledných JPEG souborů
242 je zanesena do grafu na obr. 12. Ten potvrzuje, že artefakty se zvyšující
243 se kvalitou slábnou (ale jsou stále přítomny), zatímco velikost souboru
244 roste. Pro nastavení kvality blíží se 100% je nárůst velikosti souboru
245 již velmi strmý.

246 Jestliže se pokusíme do JPEGu uložit obraz obsahující ostré barevné
247 přechody a velké jednobarevné plochy (např. kresbu na obr. 10), je možné
248 artefakty ztrátové komprese pozorovat pouhým okem i při nastavení kva-
249 lity na více než 75%. Jak je vidět na obr. 11, rozdíly zcela nezmizí ani při
250 kvalitě 98%. Pro tento typ obrazu jsou vhodnější bezztrátové formáty,
251 jako je PNG [6].

252 Snaha o ještě efektivnější kompresi zůstává aktuální i přes neustálé
253 zvyšování rychlosti internetového připojení a kapacity úložišť, neboť zá-
254 roveň roste i rozlišení snímků z digitálních fotoaparátů. Poslední dobou
255 se objevily nové formáty ukládání obrazu, například formát HEIF, který
256 využívá algoritmy komprese obrazu z video kodeku HEVC (H.265) [7],
257 nebo WebP [8], který je zase spjatý s video kodekem VP9. I v nich však
258 zůstává převod obrazu do frekvenční oblasti klíčovým krokem k efek-
259 tivní ztrátové kompresi a stejně jako v JPEG se obraz zpracovává po
260 čtvercových blocích, pro které se vžil název *makrobloky*.



(a) kvalita 98% (orig.), velikost 1100 kB



(b) kvalita 5%, velikost 64 kB

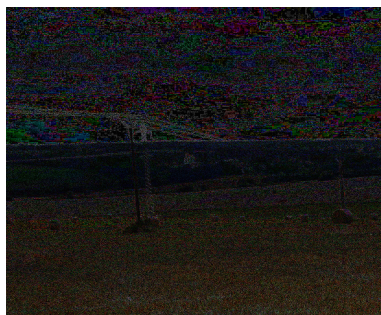
Obr. 8: Viditelná degradace fotografie venkovské krajiny při uložení JPEG s velmi nízkou kvalitou. Jedná se o výřez z fotografie na obr. 7a.

261 Závěr

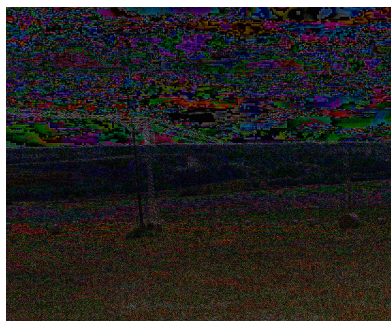
262 Matematické metody převodu dat do frekvenční oblasti založené na
263 Fourierově transformaci a dalších příbuzných transformacích mají nezastupitelné místo nejen při zpracování obrazu a při ztrátové kompresi
264 obrazu, videa a zvuku, ale též v telekomunikacích, při rekonstrukci tří-
265 rozměrných dat z magnetické rezonance nebo v rentgenové a elektronové
266 krystalografii. Svoje využití nalézají i v samotné matematice, například
267 při hledání řešení parciálních diferenciálních rovnic. Cílem našeho článku
268 bylo detailně představit jednu konkrétní metodu, tedy DFT, a to i zvídavým
269 čtenářům se znalostmi matematiky na úrovni střední školy. Snažili
270 jsme se proto vést výklad od základů a důsledně se vyhýbat konceptům
271 matematické analýzy, jako jsou limity, nekonečné řady a integrály.
272 Kdo se však těchto pojmů nebojí a zajímají ho širší souvislosti, může
273 si prostudovat např. zdroje [9, 10], které jsou dostupné online. Kromě
274 teoretického výkladu v článku nechybí ani vizuálně atraktivní praktické
275 ukázky použití DFT a demonstrace následků ztrátové komprese obrazu.
276



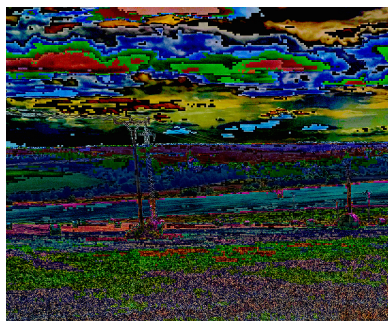
(a) kvalita 85%



(b) kvalita 50%



(c) kvalita 25%



(d) kvalita 5%

Obr. 9: Zvýrazněný rozdíl originálního obrazu a obrazu uloženého se ztrátovou kompresí JPEG pro různá nastavení kvality. Originální fotografie je na obr. 8a.

Ahoj, toto je test JPEG!

A toto je test s tentým písemem...

(a) Originál (nekomprimovaný)

Ahoj, toto je test JPEG!

A toto je test s tentým písemem...

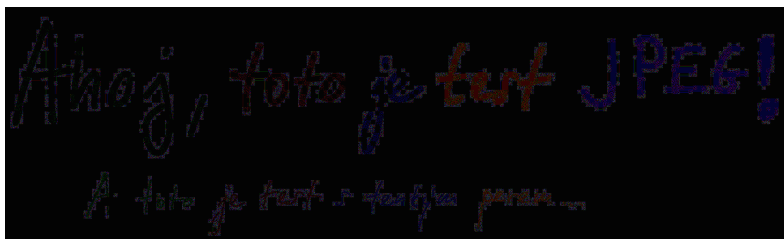
(b) kvalita 50%

Ahoj, toto je test JPEG!

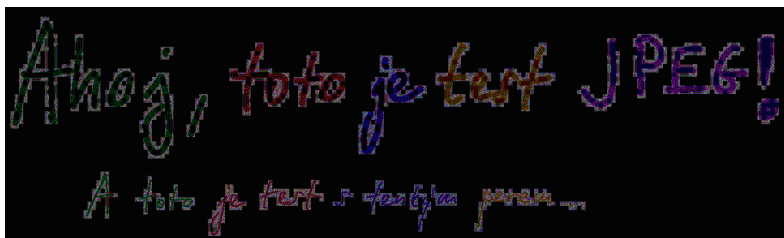
A toto je test s tentým písemem...

(c) kvalita 10%

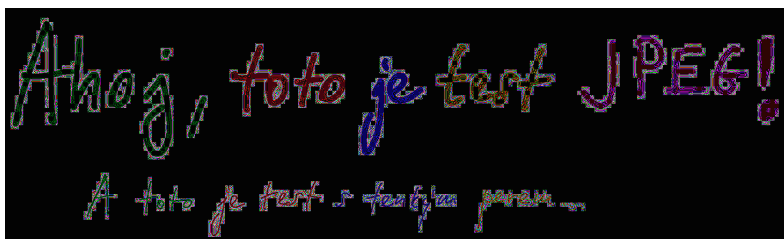
Obr. 10: Degradace obrazu s ostrými barevnými přechody při uložení do formátu JPEG.



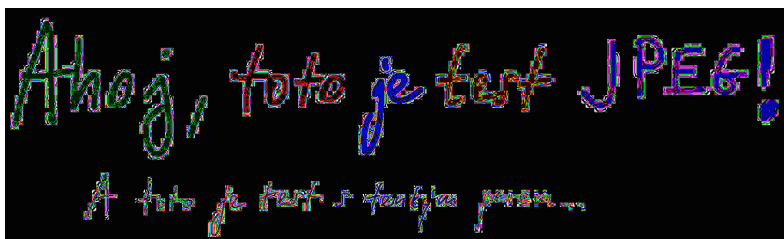
(a) kvalita 98%



(b) kvalita 85%

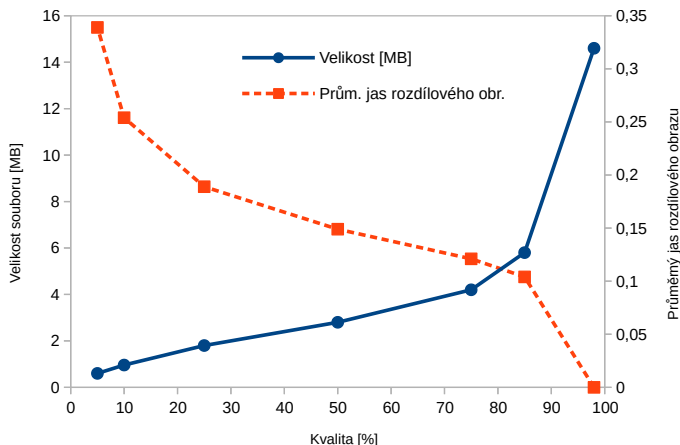


(c) kvalita 50%



(d) kvalita 10%

Obr. 11: Zvýrazněný rozdíl originálního obrazu a obrazu uloženého se ztrátovou kompresí JPEG pro různá nastavení kvality. Originální obraz s ostrými hranami je na obr. 10a.



Obr. 12: Závislost výskytu artefaktů ztrátové komprese (vyjádřeného průměrným jasnem rozdílového obrazu – viz obr. 9) a velikosti výsledného souboru na nastavení kvality JPEG. Referenčním obrazem je fotografie na obr. 7a, stažená z fotoaparátu ve formátu JPEG s kvalitou 98%.

277 Literatura

- 278 [1] W. B. Pennebaker, J. L. Mitchell: *JPEG Still Image Data Compression Standard*, Springer New York, 1993. ISBN: 978-0-442-01272-4.
279
- 280 [2] S. Hacker, *MP3: The Definitive Guide*. O'Reilly Media, 2000. ISBN: 9781565926615.
281
- 282 [3] C. Wootton: *A Practical Guide to Video and Audio Compression: From Sprockets and Rasters to Macro Blocks*, Routledge, 2005. ISBN: 9780240806303.
283
284
- 285 [4] E. Zolfagharifard: *Who do YOU see: Albert or Marilyn? Optical illusion could reveal if you need glasses*. Daily Mail Online, 3. dubna 2015. Dostupné na <https://www.dailymail.co.uk/sciencetech/article-3024005/Do-Albert-Einstein-Marilyn-Monroe-Hybrid-optical-illusion-reveal-need-wear-glasses.html>,
286
287
288
289
290 citováno 16.8.2022.
- 291 [5] P. Tišnovský: *Programujeme JPEG: diskrétní kosinová transformace (DCT)*. ROOT.cz, 4. ledna 2007. Dostupné na
292

- 293 [https://www.root.cz/clanky/programujeme-jpeg-diskretni-
kosinova-transformace-dct/](https://www.root.cz/clanky/programujeme-jpeg-diskretni-
294 kosinova-transformace-dct/), citováno 16.8.2022.
- 295 [6] G. Roelofs: *PNG - The Definitive Guide*. O'Reilly, 1999.
- 296 [7] Elemental Technologies: *HEVC Demystified - A Primer on the H.265
297 Video Codec*. The Broadcast Bridge, 2014.
- 298 [8] Google: *An Image Format for the Web*. Dostupné na [https://
developers.google.com/speed/webp](https://
299 developers.google.com/speed/webp), citováno 16.8.2022.
- 300 [9] D. Horák: *Diskrétní transformace*. Učební text ZČÚ v Plzni, 2012.
301 Dostupné na [https://mi21.vsb.cz/sites/mi21.vsb.cz/files/
unit/diskretni_transformace.pdf](https://mi21.vsb.cz/sites/mi21.vsb.cz/files/
302 unit/diskretni_transformace.pdf), citováno 16.8.2022.
- 303 [10] V. Hlaváč: Fourierova transformace v 1D a 2D. Přednáška
304 ČIRK ČVUT, dostupné na [https://people.ciirc.cvut.cz/
~hlavac/TeachPresCz/11DigZprObr/12FourierTxCz.pdf](https://people.ciirc.cvut.cz/
305 ~hlavac/TeachPresCz/11DigZprObr/12FourierTxCz.pdf), citováno
306 16.8.2022.

307 **Dodatek - diskrétní cosinová transformace**

308 Pro čtenáře, kterým se zalíbila manipulace se sumami v kapitole o
 309 DFT, nabídneme jako bonus i odvození vzorce pro DCT, která je na
 310 DFT založená. V kompresi JPEG je použita dvourozměrná DCT typu
 311 II [1]. My odvodíme tvar příslušné jednorozměrné transformace, od níž
 312 lze ke dvojrozměrné variantě přejít stejným postupem jako v (12).

313 Z původní posloupnosti f_0, \dots, f_{N-1} připravíme celkem $4N$ hodnot
 314 x_0, \dots, x_{4N-1} tak, že za ni nejprve připojíme stejnou posloupnost v opač-
 315 ném pořadí, tj. $f_{N-1}, f_{N-2}, \dots, f_1, f_0$, a následně před každou původní
 316 hodnotu vložíme nulu, aby byly nenulové hodnoty pouze na lichých
 317 indexech. Výsledná posloupnost splňuje $x_{2k} = 0, x_{2k+1} = f_k$ pro $k =$
 318 $0, 1, \dots, N-1, x_{2N} = 0$ a

$$x_\ell = x_{4N-\ell} \text{ pro } \ell = 2N+1, 2N+2, \dots, 4N-1. \quad (13)$$

Na tyto hodnoty nyní aplikujeme DFT a postupnými úpravami dostane-
 neme

$$\begin{aligned} X_n &= \sum_{k=0}^{4N-1} x_k e^{-2\pi i k n / (4N)} \stackrel{\substack{\text{jen přes} \\ \text{liché}}}{=} \sum_{k=0}^{2N-1} x_{2k+1} e^{-2\pi i (2k+1)n / (4N)} \\ &\stackrel{\substack{\text{úprava} \\ \text{exponentu}}}{=} \sum_{k=0}^{2N-1} x_{2k+1} e^{-\pi i (k + \frac{1}{2})n / N} \\ &= \sum_{k=0}^{N-1} x_{2k+1} e^{-\pi i (k + \frac{1}{2})n / N} + \sum_{k=N}^{2N-1} x_{2k+1} e^{-\pi i (k + \frac{1}{2})n / N} \\ &\stackrel{\substack{\text{první} \\ \text{a druhá} \\ \text{polovina} \\ \text{zvlášť}}}{=} \sum_{k=0}^{N-1} x_{2k+1} e^{-\pi i (k + \frac{1}{2})n / N} + \sum_{k=0}^{N-1} \underbrace{x_{2(2N-1-k)+1}}_{=x_{4N-(2k+1)}} e^{-\pi i ((2N-1-k) + \frac{1}{2})n / N} \\ &\stackrel{\substack{\text{druhá} \\ \text{suma} \\ \text{pozpátku}}}{=} \sum_{k=0}^{N-1} x_{2k+1} \left[e^{-\pi i (k + \frac{1}{2})n / N} + e^{-\pi i ((2N-1-k) + \frac{1}{2})n / N} \right] \\ &\stackrel{\substack{\text{využití (13)} \\ \text{a sloučení}}}{=} \sum_{k=0}^{N-1} x_{2k+1} \left[e^{-\pi i (k + \frac{1}{2})n / N} + e^{-2\pi i n} e^{+\pi i (k + \frac{1}{2})n / N} \right] \\ &\stackrel{\substack{\text{úprava} \\ \text{a def. } x_{2k+1}}}{=} \sum_{k=0}^{N-1} f_k \left[e^{-\pi i (k + \frac{1}{2})n / N} + \underbrace{e^{-2\pi i n}}_{=1} e^{+\pi i (k + \frac{1}{2})n / N} \right] \\ &= 2 \sum_{k=0}^{N-1} f_k \cos \left(\left(k + \frac{1}{2} \right) \frac{\pi n}{N} \right) = 2 \sum_{k=0}^{N-1} f_k \cos \left(\frac{(2k+1)\pi n}{2N} \right) \end{aligned} \quad (14)$$

319 pro $n = 0, 1, \dots, 4N - 1$. Protože X_n jsou zřejmě reálná čísla a zároveň
320 platí (11), dostáváme

$$X_{4N-n} = X_n,$$

321 což umožňuje dopočítat $X_{2N+1}, \dots, X_{4N-1}$ z X_1, \dots, X_{2N-1} . Navíc vy-
322 užítím součtového vzorce

$$\cos(\alpha - \beta) = \cos \alpha \cos \beta + \sin \alpha \sin \beta$$

323 lze snadno ukázat, že

$$X_{2N-n} = -X_n, \tag{15}$$

324 což umožňuje získat X_{N+1}, \dots, X_{2N} z X_0, \dots, X_{N-1} . Nakonec z (15)
325 plyne ještě $X_N = 0$. Celkem tedy z (14) získáváme nezávislé hodnoty X_n
326 pouze pro $n = 0, 1, \dots, N-1$. Za výsledek DCT posloupnosti f_0, \dots, f_{N-1}
327 se proto považuje posloupnost N čísel X_0, \dots, X_{N-1} .