

KAPITOLA 1

ZÁKLADNÉ INFORMÁCIE

Kompresia údajov je proces kódovania dát, pri ktorom dochádza k redukcii ich množstva. Kompresia dát je potrebná pre zvýšenie efektívnosti číslcového spracovania, prenosu a záznamu údajov (prípadne signálov vyjadrených vo forme dát). Kompresia sa dosahuje redukciou *irelevancie* (nepodstatnej zložky) a *redundancie* (nadbytočnej zložky) [18].

Redukcia nepodstatnej zložky je v procese kódovania nevratný proces. Preto dochádza k informačnej strate. Kritériom použitého stupňa redukcie nepodstatnej zložky sú vlastnosti prijímača, prípadne nároky používateľa.

Redukcia nadbytočnej zložky je vratný proces. Nadbytočnú zložku je možné v prijímači obnoviť. Tento proces preto nevedie k informačnej strate.

Metódy kódovania potom na základe použitého spôsobu redukcie môžeme rozdeliť na:

1. *reverzibilné* - bez degradácie (častejšie nazývané kompakcia),
2. *ireverzibilné* - s degradáciou (častejšie nazývané kompresia) [18].

Pre vyjadrenie stupňa kompresie je definovaný *kompresný pomer* [39]

$$K = \frac{c_1}{c_2}, \quad (1.1)$$

kde c_1 je množstvo dát pred kompresiou a c_2 je množstvo dát po kompresii

alebo tiež *relatívna redundancia* dát (ktorá sa zhoduje s irelevanciou pri metódach kódovania so stratou) [39]

$$R = 1 - \frac{1}{K}. \quad (1.2)$$

Pri kompresii digitálneho obrazu uvažujeme tri typy redundancie dát: *kódová* redundancia, *medzisymbolová* (príp. medzisnímková, medzibloková) redundancia a *psychovizuálna* redundancia [39].

1.1 KÓDOVÁ REDUNDANCIA

Predpokladajme, že diskretná náhodná veličina r_i nadobúdajúca hodnoty z intervalu $\langle 0, 1 \rangle$, reprezentuje úroveň (intenzitu) jasnosti obrazu. Nech sa hodnota r_i vyskytuje s pravdepodobnosťou $p_i(r_i)$, ktorú definujeme ako relatívnu početnosť.

Potom [39]

$$p_i(r_i) = \frac{h_i}{N_1 \cdot N_2},$$

$$i = 0, 1, 2, \dots, L-1, \quad (1.3)$$

kde L je počet úrovní jasu, h_i je počet výskytu i -tej úrovne jasu a N_1 a N_2 sú rozmery obrazu.

Počet bitov potrebných na reprezentáciu každej hodnoty r_i , nech je $l(r_i)$. Potom priemerný počet bitov potrebných na reprezentáciu každého bodu obrazu je [39]

$$L_b = \sum_{i=0}^{L-1} l(r_i) \cdot p_i(r_i). \quad (1.4)$$

Celkový počet bitov potrebných na zakódovanie daného obrazu potom je $N_1 \cdot N_2 \cdot L_b$.

Príklad

Máme 8-úrovňový obraz s relatívnou početnosťou úrovní: 0.18, 0.26, 0.20, 0.17, 0.08, 0.05, 0.03, 0.03. Pre zakódovanie binárnym kódom potrebujeme 3 bity pre jednu úroveň šedej. Potom L_b je 3 bity. Navrhujeme iný kód, ktorý napr. odhadom reprezentuje úrovne s malou početnosťou väčším počtom bitov, než úrovne s väčšou početnosťou. Použijeme napr. kód s nerovnakou dĺžkou slova postupne pre uvedené početnosti: 11, 01, 10, 001, 0001, 00001, 000001, 000000.

$$L_b = \sum_{i=0}^{L-1} l(r_i) \cdot p(r_i) = 2 \cdot (0,18) + 2 \cdot (0,26) + 2 \cdot (0,20) + 3 \cdot (0,17) + 4 \cdot (0,08) + 5 \cdot (0,005) + 6 \cdot (0,03) + 6 \cdot (0,03)$$

$$= 2,7 \text{ bitu}$$

$$K = 1, 11$$

To znamená, že binárny kód je oproti nami navrhnutému kódu redundantný.

V príklade ukázaný proces kódovania sa nazýva kódovanie s variabilnou dĺžkou slova. Neskôr si ukážeme niektoré spôsoby vytvárania kódov s variabilnou dĺžkou slova. Ako vidno z príkladu, dôležitou veličinou pri ich konštrukcii bude početnosť výskytu jednotlivých úrovní, či čísel.

1.2 MEDZISYMBOLOVÁ REDUNDANCIA

Väčšina reálnych dát z praxe, zvlášť dáta obrazové, je štatisticky veľmi závislých. Závislosť (korelovanosť) dát sa prejavuje vo všetkých smeroch. Preto hovoríme o korelácii riadkovej, stĺpcovej, ale aj krížovej. Znamená to, že údaje sa správajú podľa nejakej štatistickej závislosti (rozdelenia). Takéto dáta sú potom veľmi závislé (prípadne podobné) dátam zo svojho okolia. Vtedy hovoríme o *medzibodovej* redundancii [39].

Existuje však veľa iných závislostí. Ak obraz znázorňuje skupinu rovnakých predmetov, je predpoklad, že niektoré časti obrazu budú veľmi podobné, ak nie úplne rovnaké. Vtedy hovoríme o *geometrickej* redundancii [39].

V pohyblivom obraze alebo v obraze priestorovom je veľká závislosť bodov v treťom rozmere, teda medzi jednotlivými snímkami alebo rezmi. V takomto prípade pripadá do úvahy *medzisnímková* redundancia [39].

Podobné úvahy platia i pre kódovanie so stratou informácie. V týchto kódovacích metódach sa tento princíp veľmi využíva. Pretože korelácia dát nám vlastne priamo hovorí o pomernej podobnosti

symbolov, snahou je vypracovávať také postupy, aby z dát, ktoré sú korelované, boli vytvorené dáta nekorelované. Tieto by mali dostatočne opísať príslušný obraz.

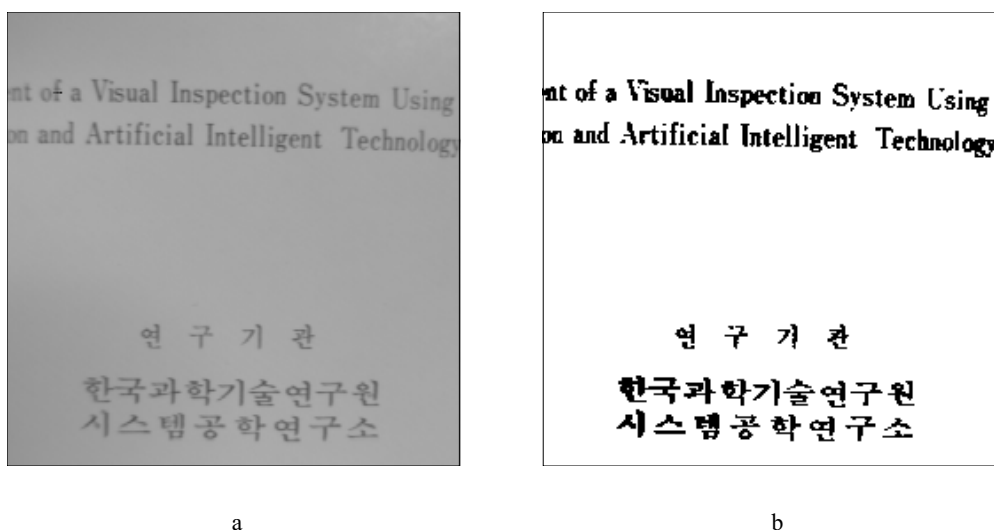
1.3 PSYCHOVIZUÁLNA REDUNDANCIA

Psychovizuálnu redundanciu si vysvetlíme na jednoduchom príklade. Vieme, že ľudské oko je schopné rozoznávať 256 šedých úrovní lineárne kvantovaných. Preto preň bude napr. zbytočné zariadenie, ktoré buď sníma alebo zobrazuje viac úrovní, než je 256. Je to nad schopnosti ľudského vnímania, preto aj takýto snímok, určený pre človeka, je často zbytočne presný. To znamená nadbytočný alebo nepodstatný. Vtedy hovoríme o zbytočnej presnosti kvantizácie obrazu. Neplatí to všeobecne. Napríklad v počítačovej tomografii je nevyhnutné využiť všetkých 12 bitov, lebo nesú dôležitú informáciu o rôznych štruktúrach. Prakticky sa dá samozrejme využiť iba 8 bitov, čo však znamená prejsť na výber potrebných okien v pôvodnej šírke údajov. V prípade kvantovania existuje aj iná úvaha. Niektoré snímky majú svetlosť len v malom rozsahu. Preto v tomto rozsahu je vhodnejšie presne kvantovať a v rozsahu malého výskytu úrovní kvantovať s menšou presnosťou. Vtedy sa nám aj s menším počtom kvantizačných úrovní podarí zachovať obraz v dostatočnej kvalite. Tak ako v jednorozmernom spracovaní tieto kvantizátory budeme nazývať nelineárnymi.

Podobný príklad je aj s priestorovou presnosťou snímania, teda s hustotou vzorkovania. Hustota vzorkovania závisí nielen od snímacieho a zobrazovacieho zariadenia, ale je podmienená aj vizuálnymi schopnosťami používateľa a aplikáciou, pre ktorú bol obraz snímaný. Iný počet bodov na opis obrazu potrebuje používateľ videotelefónu a iný počet (hustotu) si žiada sledovanie televízie.

Ďalší z príkladov psychovizuálnej redundancie je sledovanie pohyblivého obrazu. V mnohých aplikáciách, napr. videotelefónnej službe, je pre používateľa podstatný objekt, ale nie jeho pozadie. Vtedy na opis pozadia nie je potrebné uchovávať všetko, čo nám snímacie zariadenie núka. Pre používateľa je potrebné presnejšie opísať pohybujúci sa objekt.

Z praktického života vieme, že napr. pre počuteľnosť určitých zvukov je zbytočné prenášať a spracovávať celý rozsah pásma počuteľnosti ľudského ucha. Ide napr. o telefónny hovor, meracie signály, ale aj signály biomedicínskeho charakteru. Podobný prípad je aj pri prenose a uchovávaní digitálneho obrazu. Pre opis obrazov s jemnou štruktúrou potrebujeme samozrejme širšie frekvenčné pásmo, než pre obrazy s malými zmenami. Toto v ďalšom využijeme na konštrukciu kódov, ktoré nebudú obraz opisovať v časovej, ale v spektrálnej oblasti. Príklady psychovizuálne redundandných a upravených obrazov sú na [obr. 1.1](#) a [obr. 1.2](#).



Obr. 1.1 Obraz s 256 úrovňami jasu (a), ten istý obraz upravený na 2 úrovne jasu (b)



Obr. 1.2 Obráz Claudia 256 x 140 obrazových bodov s 256 úrovňami jasu (a), ten istý obraz kvantovaný lineárne na 16 úrovni jasu (b), nelineárne (c)

1.4 KRITÉRIÁ KVALITY

Pri kompresii dát metódami, ktoré spôsobujú ich konečnú degradáciu, je potrebné stanoviť kritériá pre určenie kvality rekonštruovaného (dekódovaného) obrazu. Definujú sa objektívne a *subjektívne* kritériá. Pre objektívne kritériá je treba definovať najprv originál a rekonštruovaný obraz. Nech body originálu sú $x(n_1, n_2)$ a prislúchajúce body rekonštruovaného obrazu sú $\hat{x}(n_1, n_2)$. Potom *odchýlka* rekonštruovaného obrazu od originálu v každom bode je [39], [19]

$$e(n_1, n_2) = \hat{x}(n_1, n_2) - x(n_1, n_2), \quad (1.5)$$

stredná kvadratická odchýlka je definovaná ako

$$e_{MSE} = \frac{1}{N_1 \cdot N_2} \sum_{n_2=0}^{N_2-1} \sum_{n_1=0}^{N_1-1} e(n_1, n_2)^2, \quad (1.6)$$

odstup signál - šum je definovaný [39], [71] vzťahom

$$SNR_{MSE} = \frac{1}{N_1 \cdot N_2 \cdot e_{MSE}} \sum_{n_2=0}^{N_2-1} \sum_{n_1=0}^{N_1-1} x(n_1, n_2)^2, \quad (1.7)$$

vrcholový odstup signál - šum, ktorý sa vzťahuje na maximálnu úroveň [71]

$$PSNR = \frac{1}{e_{MSE}} \cdot 255^2, \quad (1.8)$$

často bývajú uvádzané v decibeloch.

Pretože v množstve prípadov (hlavne keď nie je k dispozícii originál) tieto tzv. objektívne kritériá nedávajú dostatočnú predstavu o kvalite rekonštruovaného obrazu, bolo zavedených niekoľko typov subjektívneho hodnotenia výsledku [39]. Buď sa obraz hodnotí známkami $\{-3, -2, -1, 0, 1, 2, 3\}$, kde -3 predstavuje najhoršie hodnotenie, alebo sú zavedené slovné opisy: vynikajúci, pekný, dostatočný (doslovne "ujde"), na hranici, horší a nepoužiteľný. Tieto kritériá sú závislé len od používateľa, prípadne od tvorcu rôznych báz údajov. Predpokladá sa, že ide o človeka dostatočne kvalifikovaného

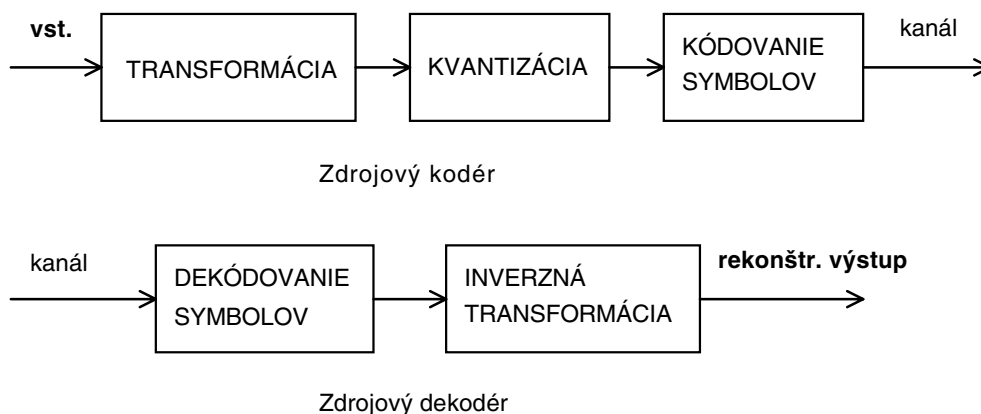
na to, aby mohol posúdiť kvalitu pre ďalších používateľov. Aj napriek tomu, že subjektívne kritériá sú v tomto prípade asi smerodajnejšie, než čísla objektívnych kritérií, často sa objektívnym kritériám nedá vyhnúť. Ide predovšetkým o kódovanie pohyblivého obrazu v reálnom čase, keď nie je samozrejme dosť času na subjektívne hodnotenie jednotlivých snímkov.

1.5 MODEL OBRAZOVÉHO KODÉRA A DEKODÉRA

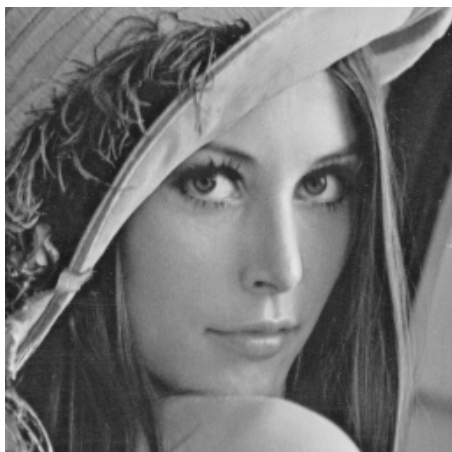
Systém pre kompresiu obrazu si môžeme znázorniť ako spojenie dvoch základných blokov: kodéra a dekodéra. Vstupné obrazové dáta sú kódované do tvaru symbolov, potrebných pre prenos, či uchovávanie. Po prenose kanálom sú údaje dekódované, resp. rekonštruované. V dekodéri je teda z prijatých symbolov viac alebo menej úspešne generovaný výstupný obraz.

Kodér sa dá pre potreby opisu kompresných postupov rozdeliť na zdrojový kodér a kanálový kodér. Kanálový kodér kóduje dáta do tvaru, potrebného pre technickú realizáciu samotného prenosu. Nás bude zaujímať zdrojový kodér, pretože ten je zodpovedný za zníženie redundancie obrazových dát. Skladá sa všeobecne [39] z troch častí. V prvej časti ide o transformáciu dát do takého tvaru, aby sa dosiahla čo najväčšia dekorelácia medzibodových štatistických závislostí. Je to vratný proces. Táto časť sa zdá byť najdôležitejšou, pretože práve ona nám môže pripraviť dáta podľa ich dôležitosti pre opis obrazu množinou štatisticky nezávislých dát. Pole týchto dát je často reálne a ich hodnoty nie vždy rovnakým dielom prispievajú k celkovému psychovizuálnemu dojmu z obrazu. Takisto nie vždy je potrebné každý údaj uchovať s rovnakou presnosťou. Preto ďalším blokom zdrojového kodéra býva kvantizátor, či už v tvare separovaného číselného zaokrúhľovania, pridelenia poradového čísla úrovne kvantovania alebo priradenia vektora, či matice jednotlivým dátovým segmentom. Tým sa dosahuje redukcia psychovizuálnej redundancie.

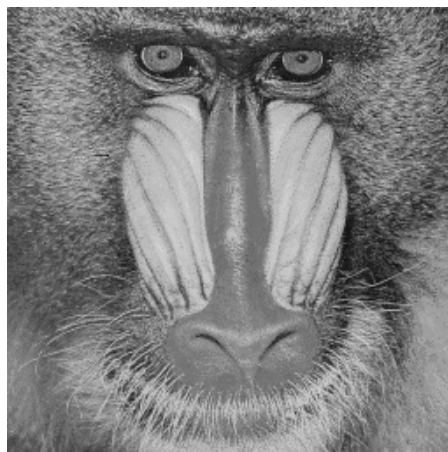
Pretože obraz sa obvyčajne spracováva po blokoch, a jednotlivé úrovne sa v blokoch opakujú s rôznou početnosťou, je vhodné, aby redundancia kvantovaných dát bola ďalej redukovaná. Tu prístupujú metódy kódovania, ktoré sú pre opakujúce sa dáta pripravené. Najčastejšie to bývajú kódy s variabilnou dĺžkou slova. Tento kodér budeme nazývať kodér symbolov, prípadne entropický kodér, ale súvislosti s entropiou si vysvetlíme až neskôr. Takto kódované dáta sú potom pripravené pre kanálový kodér, ktorý by im mal zabezpečiť dostatočnú odolnosť voči kanálovému šumu, spôsobujúcemu dátovú chybovosť. Dekodér má samozrejme opačnú funkciu ako kodér. Chýba v ňom len "dekvantizátor", pretože jeho úlohu plní už dekodér symbolov. Blokové schéma kodéra a dekodéra je na [obr.1.3](#). Na [obr. 1.4](#) sú originály obrazov, ktoré budú použité v ďalšom na ukážku účinnosti metód kompresie obrazu.



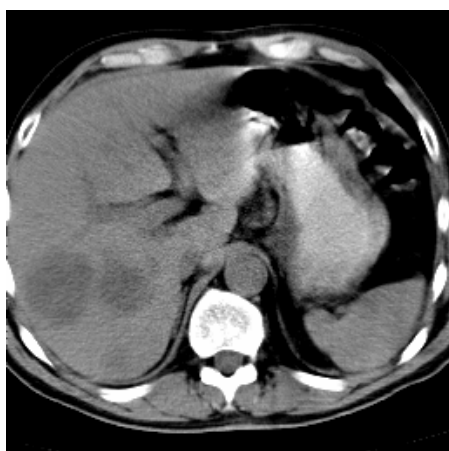
Obr. 1.3 Model zdrojového kodéra a zdrojového dekodéra



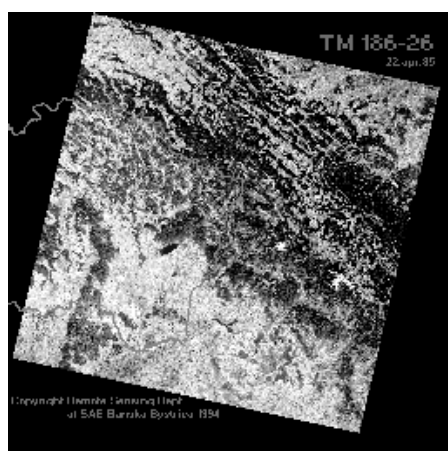
a



b



c



d

Obr. 1.4 Originály obrazov veľkosti 256 x 256 obrazových bodov s počtom úrovní jasu 256: výrez obrazu Lena (a), baboon (b), pečeň (c), aerokozmická snímka (d)