# Vpliv podvzorčenja v krominančnem prostoru na kakovost zapisa slikovnega gradiva

**Jan Pelicon[1], Urban Burnik[2]**

[1] *Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Fakulteta za elektrotehniko*
[2]*Univerza v Ljubljani, Fakulteta za elektrotehniko,*

E-pošta: *jp1467@student.uni-lj.si*

***Povzetek.*** *Pri projektu smo testirali in ovrednotili vpliv znanih in nekaj manj znanih eksperimentalnih tipov podvzorčenja v krominančnem prostoru. Kakovost podvzorčenja smo kvantitativno ovrednotili z merama strukturne podobnosti (SSIM) in vršnega razmerja signal-šum (PSNR) glede na gostoto podatkovnega zapisa v bitih na slikovni element (bits per pixel; bpp). Preizkusili smo tudi adaptivno podvzorčenje, ki na podlagi standardne deviacije makrobloka določi ustrezen tip podvzorčenja. Pri tem je potrebno podati tudi podatek o tipu podvzorčenja za posamezen blok.*

## 1 Uvod

Barvno podvzorčenje je pogosto uporabljen postopek pri kodiranju slikovnega in video materiala. Pri tem izbranim kanalom barvnega formata zmanjšamo ločljivost in s tem zmanjšamo količino podatkov za ceno nižje kvalitete zapisa. Prihranek in izguba kvalitete sta odvisna od barvnega formata in tipa podvzorčenja. Tip podvzorčenja definira, kako in koliko bo podvzorčen posamezen barvni kanal.

Krominančno podvzorčenje je definirano specifično za YCbCr digitalni barvni format [1], kjer Y kanal (ang. Luma) nosi informacijo o svetlosti, Cb (ang. Chroma blue) in Cr (ang. Chroma red) pa predstavljata razliko modre in rdeče kroma komponente. YCbCr je poznan tudi kot YPbPr ali Y'CbCr. Pogosto se ga zamenjuje z YUV formatom, ki pa se uporablja izključno pri analognem prenosu signalov. Pretvorba iz 8 bitnega RGB formata v 8 bitni YCbCr format je definirana preko naslednjih enačb [2]:

$$Y = 16 + \frac{65.738R}{256} + \frac{129.057G}{256} + \frac{25.064B}{256} \quad (1)$$

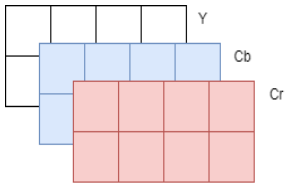
$$Cb = 128 + \frac{37.945R}{256} + \frac{74.494G}{256} + \frac{112.439B}{256} \quad (2)$$

$$Cr = 128 + \frac{112.439R}{256} + \frac{94.154G}{256} + \frac{18.285B}{256} \quad (3)$$

Cilj krominančnega podvzorčenja je zmanjšanje količine podatkov, kar dosežemo z nižanjem barvne ločljivosti Cb in Cr kanala. Pri tem upoštevamo, da je človeško oko percepcijsko bolj občutljivo na spremembe v intenziteti kot na spremembe v barvi, zato postopka v prostoru RGB ne moremo opraviti. Svetlostnega kanala Y običajno ne podvzorčimo. Podvzorčenje je neobrnljiva operacija, saj zavržene informacije ni mogoče povrniti.

## 2 Tipi krominančnega podvzorčenja

Tipi podvzorčenja [3] [4] so navadno izraženi s tridelnim razmerjem v notaciji **J:a:b**. Osnovni blok podvzorčenja običajno obsega 4×2 slikovna elementa (Slika 1). Parameter J predstavlja število vzorcev luminančnega kanala (Y) v posamezni vrstici osnovnega bloka. Vrednosti **a** in **b** se nanašata na število vzorcev obeh krominančnih kanalov (CbCr) v prvi (**a**) oziroma drugi (**b**) vrstici. Pri tem velja, da je **b** manjši ali enak **a**, saj je prva vrstica glavni vir vzorcev, druga vrstica pa doprinese enako količino vzorcev ali pa kopira vrednosti prve vrstice. Skozi čas so se uveljavili različni tipi podvzorčenja.




Slika 1. Osnovni blok za izvajanje operacije podvzorčenja.

### 2.1 Tip 4:4:4

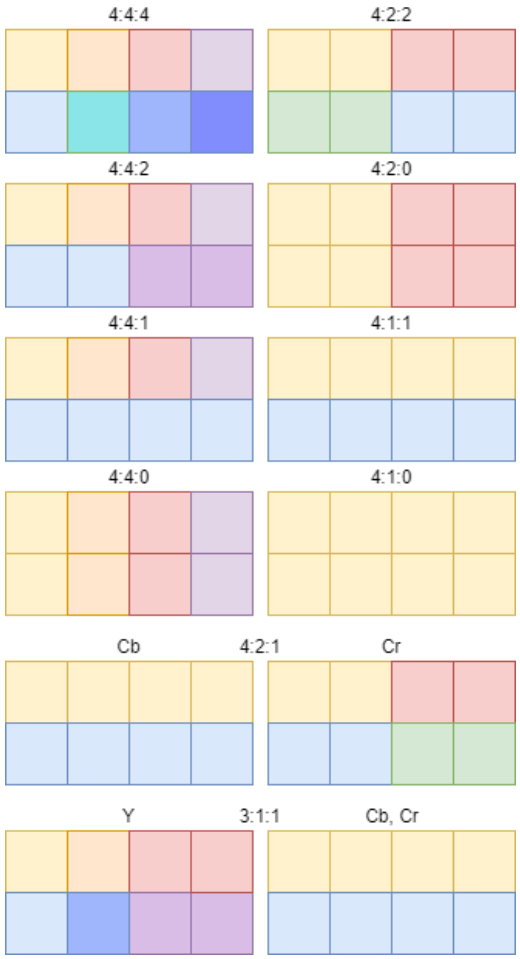
Tip 4:4:4 predstavlja sliko, ki ni podvzorčena, saj ima vseh 8 elementov osnovnega bloka lastne vrednosti vseh treh barvnih kanalov YCbCr. Število bitov na slikovni element (bpp) v primeru 8 bitov na kanal tako znaša 24 bpp. Torej lahko vsakemu od 8 slikovnih elementov določimo svojo barvo.

### 2.2 Tip 4:2:1

Tip 4:2:1 je zastarel, tolmačenje J:a:b notacije pa ni skladno s preostalimi zapisi. Pri tem tipu ima vrstica bloka kanala Cb dvakrat manjšo ločljivost od vrstice bloka kanala Cr in štiri krat manjšo ločljivost v primerjavi z vrstico Y kanala. Y kanal prispeva 8 bpp, Cr kanal 4 bpp, Cb pa 2 bpp, kar skupno znaša 14 bpp.

### 2.3 Tip 4:1:1

Pri tipu 4:1:1 je v vsaki vrstici krominančnih kanalov po en vzorec. Čeprav je barvna ločljivost tega tipa precej okrnjena, je še vedno prisoten v nekaterih formatih kot so DVCPRO (NTSC in PAL), NTSC DV in DVCAM ter D-7. Tip 4:1:1 potrebuje 12 bpp, 8 bpp za Y kanal ter 2 bpp za posamezen krominančni kanal, kar skupno znaša polovico bpp primera brez podvzorčenja.




Slika 2. Vizualni prikaz različnih tipov podvzorčenja.

### 2.4 Tip 4:2:0

Najbolj široko uporabljen in poznan tip podvzorčenja je 4:2:0. Ta tip se uporablja v različnih kodekih, standardih in formatih kot so na primer vsi ISO/IEC MPEG kodeki ter H.26x standardi. Prav tako je uporabljen pri DVD in Blu-ray Disc video formatu. Tako kot tip 4:1:1 se uporablja pri DVCAM in tudi PAL DV ter HDV (ang. High Definition Video). Prav tako je pogost pri JPEG, JFIF in MJPEG implementacijah. Kot pri tipu 4:1:1 je količina krominančnih podatkov zmanjšana za faktor 4, kar v tem primeru predstavlja enakomerneje razporejeno polovično vertikalno in horizontalno ločljivost.

### 2.5 Tip 4:1:0

Uporaba podvzorčnega tipa 4:1:0 je redka, zasledimo ga lahko le nekaterih kodekih. Pri tem tipu je značilno da imata krominančna kanala Cb in Cr v osnovnem bloku le en vzorec, kar prinaša večje prihranke za ceno slabše kvalitete. Obstajajo tudi izboljšave kot so 4:1:0.5 ali 4:1:0.25, da dosežemo višjo kvaliteto kot pri osnovnem 4:1:0. Število bpp je pri tem tipu enako 10. Kanal Y potrebuje 8 bpp, za posamezen krominančni kanal pa samo 1 bpp.

### 2.6 Tip 3:1:1

Tip 3:1:1 ni pogost. Uporabljen je le pri nekaterih HD kamerah znamke Sony. Tudi ta tip nekoliko odstopa od standardne notacije **J:a:b**, saj J v tem primeru označuje zmanjšanje frekvence vzorčenja luminančnega kanala Y. V tem primeru je to ¾ vzorcev v vrstici. Konkretno v primeru HD ločljivosti to predstavlja 1440 vzorcev na vrstico siceršnje velikosti 1920. Preostali vrednosti a in b specificirata, da ima navedeni tip popolno vertikalno in četrtinsko horizontalno krominančno ločljivost.

### 2.7 Tip 4:2:2

Preostane nam še tip 4:2:2. V uporabi je pri digitalizaciji kompozitnega analognega video gradiva, pri katerem zaradi nižje pasovne širine analogne krominančne komponente v horizontalni smeri podvzorčenje izvedemo, v vertikalni smeri pa ohranimo vso razpoložljivo kakovost zapisa. Za razliko od tipa 4:2:0, ima ta tip v tudi drugi vrstici po dva vzorca. Število bpp za posamezen krominančni kanal je 4, kar znaša skupno 16 bpp.

### 2.8 Eksperimentalni tipi 4:4:2, 4:4:1, 4:4:0, 4:2:2

Pri projektu smo se odločili testirati tudi tipe, ki trenutno nimajo praktične uporabe oziroma niso uporabljeni v nobenem standardu. Namen tega je, da jih ovrednotimo v primerjavi z ostalimi standardnimi tipi in tako ugotovimo njihov potencial oziroma razlog za njihovo neuporabo.

Značilne lastnosti tipa 4:4:2 so polna ločljivost prve vrstice osnovnega bloka in polovična ločljivost druge vrstice. Kanal Y ni podvzorčen in zanj potrebujemo 8 bpp, za posamezen krominančni kanal pa 6 bpp, kar skupno znaša 20 bpp.

Tip 4:4:1 se od tipa 4:4:2 razlikuje le v ločljivosti druge vrstice. Posamezen krominančni kanal tega tipa opišemo s 5 bpp, kar skupno znaša 18 bpp.

Tip 4:4:0 je prav tako soroden zgoraj opisanima tipoma. Njegova značilnost je, da druga vrstica osnovnega bloka nima lastnih vzorcev, temveč samo kopira vzorce prve vrstice. Ta lastnost je značilna tudi za pogosto uporabljen tip 4:2:0. Posamezen krominančni kanal tega tipa opišemo s 4 bpp, kar skupno znaša 16 bpp.

## 3 Orodja in postopki

Analizo vpliva podvzorčenja in vrednotenje smo opravili v jeziku Python. Za branje, prikazovanje in pretvarjanje barvnih formatov slikovnega materiala uporabljamo odprtokodno knjižnico OpenCV, ki za hranjenje slikovnih podatkov uporablja knjižnico numpy, ki vsebuje operacije za delo s tabelami in matrikami.

Za vrednotenje smo uporabili dve metriki, ki zgoščen slikovni material ovrednotita z dveh različnih vidikov – prva temelji na primerjavi vrednosti istoležnih slikovnih elementov, druga pa se bolje približa človeški vizualni percepciji.

Za ustrezno ovrednotenje je potrebno uporabiti nezgoščeno slikovno zbirko. Za potrebe računalniškega vida in testiranje zgoščevalnih algoritmov za to obstaja že uveljavljena podatkovna zbirka Kodak, ki vsebuje 25 nezgoščenih slik, katero smo uporabili tudi pri tem projektu.

### 3.1 PSNR in MSE

Metrika PSNR (ang. Peak Signal to Noise Ratio) je definirana preko metrike MSE (ang. Mean Squared Error), ki predstavlja povprečje kvadratov razlik istoležnih vrednosti in je definirana z enačbo 4.

$$MSE(x,y) = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(x(n,m) - y(n,m)\right)^2 \quad (4)$$

Težava, ki jo je mogoče zaznati pri metriki MSE, je ta, da je ta odvisna od razpona vrednosti slikovnih elementov. V primeru 8 bitov je to območje od 0 do 255. Metrika PSNR se tej težavi izogne tako, da pri izračunu upošteva tudi barvno globino slikovnega gradiva. Definirana je kot razmerje med največjo močjo signala in močjo šuma (napake), ki se pojavlja v signalu, kot izraža enačba 5. Signal v tem primeru predstavljajo vrednosti slikovnega gradiva. Izraža se z logaritemsko skalo v decibelih (dB).

$$PSNR = 10\, log_{10}\left(\frac{MAX}{MSE^2}\right) \quad (5)$$

Vrednost MAX predstavlja razliko med največjo in najmanjšo vrednostjo, ki je v primeru 8 bitov enaka 255. V primeru, da sta signala enaka, šum ni prisoten, vrednost PSNR pa je v tem primeru enaka ∞.

### 3.2 SSIM

Metrika SSIM (ang. Structural Similarity Index Measure) se uporablja za oceno podobnosti med dvema slikama, njena vrednost pa ne temelji na primerjavi istoležnih vrednosti, temveč na primerjavi različnih lastnosti in strukture slik, ki ju primerjamo. Lahko rečemo, da vrednost poskuša zajeti človekovo subjektivno percepcijo kvalitete slike. Metrika izhaja iz ideje, da so piksli pogosto medsebojno odvisni, če so si prostorsko blizu. Te odvisnosti vsebujejo pomembne informacije o objektih in strukturi slike, torej lahko preko njih ovrednotimo podobnost med slikama. Izračun podobnosti SSIM poteka preko drsečih oken $x$ in $y$, ki potujeta preko istoležnih območij obeh slik in pridobita vrednost SSIM za dano območje. Formula je sestavljena iz treh komponent $l$, $c$ in $s$, predstavljenih v enačbah 6, 7 in 8. V komponentah nastopajo konstante $C_1$, $C_2$ in $C_3$. Komponenta $l$, kjer je $\mu$ pričakovana vrednost, primerja intenziteto med opazovanima območjema.

$$l(x,y) = \frac{2\mu_x\mu_y + C_1 xy}{\mu_x^2 + \mu_y^2 + C_1} \quad (6)$$

Komponenta c, kjer je $\sigma$ standardni odklon, primerja kontrast med opazovanima območjema.

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (7)$$

Ostane še komponenta $s$, kjer je $\sigma_{xy}$ kovarianca med območjema $x$ in $y$ in poskuša zajeti strukturno podobnost.

$$s(x,y) = \frac{2\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (8)$$

$C_1$, $C_2$ in $C_3$, ki so definirane v enačbah 9, 10 in 11, so odvisne od $L$, ki predstavlja razliko med največjo in najmanjšo vrednostjo, ki se pojavi v sliki.

$$C_1 = (0{,}01 \times L)^2 \quad (9)$$

$$C_2 = (0{,}03 \times L)^2 \quad (10)$$

$$C_3 = \frac{C_2}{2} \quad (11)$$

Definirali smo vse tri komponente za izračun SSIM, ki se izračuna kot uteženo kombinacijo teh. Izračun prikazuje enačba 12.

$$SSIM(x,y) = l(x,y)^{\propto} \times c(x,y)^{\beta} \times s(x,y)^{\gamma} \quad (12)$$

Z vrednostmi $\propto, \beta, \gamma$ je mogoče komponente poljubno utežiti. Njihova privzeta vrednost je enaka 1.

## 4 Rezultati

Zgradili smo preprost cevovod za izvajanje krominančnega podvzorčenja (izvorna koda je dostopna na https://github.com/JanPelicon/OMMV-Projektna-Naloga).

Slikovni material najprej pretvorimo iz RGB v YCbCr barvni format. Nato sledi prirejanje velikosti slike, da se ta ujema z večkratnikom širine in dolžine osnovnega bloka slikovnih elementov.

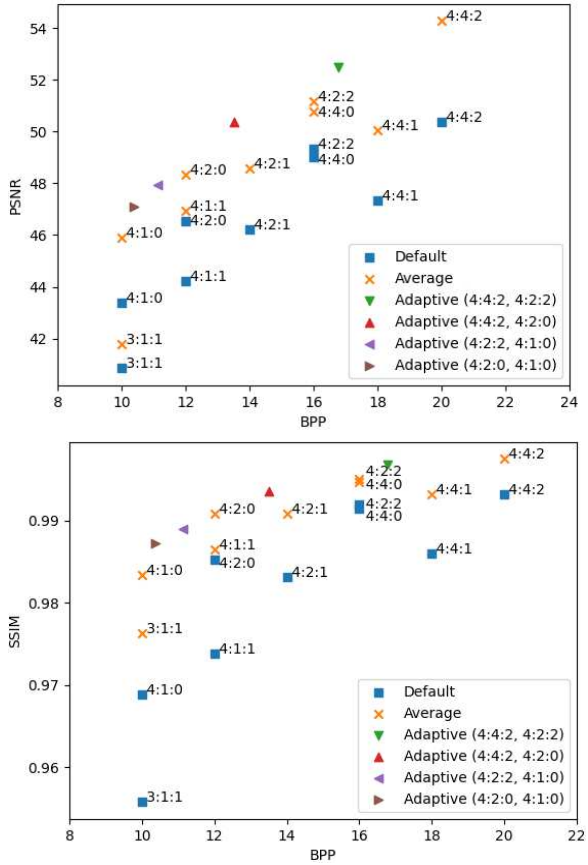
Sliko razdelimo na osnovne bloke, nad katerimi izvedemo operacijo krominančnega podvzorčenja. Za pridobitev barvnega vzorca smo uporabili dve različni tehniki. Po privzeti tehniki smo za vzorec uporabili prvo vrednost v vrsti vrednosti, katere bodo imele isti vzorec. Po drugi tehniki smo izračunali povprečje teh vrednosti.

Preizkusili smo tudi adaptivno krominančno podvzorčenje, ki za makro blok velikosti 16x16 izračuna standardno deviacijo bloka in na podlagi te določi, kateri tip podvzorčenja bo uporabljen. Pri implementaciji smo se odločili uporabiti različne kombinacije dveh različnih tipov podvzorčenja. Omejili smo se na 4 tipe podvzorčenja, ki so imeli najboljše razmerje med kvaliteto in količino bitov potrebnih za zapis. To so tip 4:4:2, 4:2:2, 4:2:0 in 4:1:0. Meja za prehod med tipoma je bila standardna deviacija bloka z vrednostjo 4.0. Če je bila standardna deviacijo večja od te vrednosti, smo uporabili tip podvzorčenja, ki ponudi boljšo kvaliteto, saj

so vrednosti v bloku bolj raznolike in jih je smiselno natančneje vzorčiti.

Rezultati opisanih tipov podvzorčenja z različnimi tehnikami so prikazani v tabeli 1 ter na sliki 3. Kot lahko vidimo, se tehnika s povprečenjem vrednosti za pridobitev vzorca bolje obnese pri vseh tipih podvzorčenja. V povprečju je opaziti občutno razliko med tehnikama, pri PSNR to znaša cca. 2 dB, pri SSIM pa je razlika vrednosti v območju od 0.003 do 0.02.

Adaptivno podvzorčenje prinese manjšo izboljšavo za ceno dodatne kompleksnosti, saj izkorišča prednosti obeh uporabljenih tipov podvzorčenja. Pri tem upoštevamo, da je potrebno za vsak makro blok 16x16 dodati 1 bit, ki nosi informacijo o tipu podvzorčenja makrobloka.




Slika 3. Grafa, ki prikazujeta razmerja kvalitete (PSNR, SSIM) in števila potrebnih bitov na slikovni element (bpp).

## 5 Zaključek

Uspešno smo ovrednotili obstoječe in nekatere eksperimentalne tipe podvzorčenja v krominančnem prostoru. Izkazalo se je, da je način izbire vzorca pomemben za doseganje boljše kvalitete. Poleg tega je iz rezultatov razvidno, da se tipi podvzorčenja pri isti vrednosti bpp lahko precej razlikujejo po doseženi kvaliteti. Opazno je, da je smiselno uporabiti enakomerno vzorčenje po horizontali in vertikali (tip 4:2:2 proti 4:4:0). Če imamo redko horizontalno vzorčenje, se to pozna na končni kvaliteti.

Potrebno bi bilo podrobneje ovrednotiti adaptivno podvzorčenje z različnimi mejnimi vrednostmi za prehod med tipoma. Prav tako bi bilo smiselno sistem nadgraditi z večjim številom tipov za posamezen makro blok.

Smiselno bi bilo preizkusiti tudi podvzorčenja na različno velikih osnovnih blokih in temu primernih tipih podvzorčenja [5]. Pri temu projektu smo se omejili na standarden 4x2 blok, ki pa ni nujno najbolj učinkovit oziroma ne ponuja dovolj raznolikosti.

Tabela 1. Rezultati testiranja obeh tehnik vzorčenja in adaptivnega podvzorčenja.

| | bpp | SSIM | PSNR |
|---|---|---|---|
| 4:4:4 | 24 | 1.000 | ∞ |
| 4:4:2 default | 20 | 0.993 | 50.372 |
| 4:4:1 default | 18 | 0.986 | 47.335 |
| 4:4:0 default | 16 | 0.991 | 49.030 |
| 4:2:2 default | 16 | 0.992 | 49.328 |
| 4:2:1 default | 14 | 0.983 | 46.206 |
| 4:2:0 default | 12 | 0.985 | 46.524 |
| 4:1:1 default | 12 | 0.974 | 44.226 |
| 4:1:0 default | 10 | 0.969 | 43.394 |
| 3:1:1 default | 10 | 0.956 | 40.854 |
| 4:4:2 average | 20 | 0.998 | 54.267 |
| 4:4:1 average | 18 | 0.993 | 50.054 |
| 4:4:0 average | 16 | 0.995 | 50.781 |
| 4:2:2 average | 16 | 0.995 | 51.159 |
| 4:2:1 average | 14 | 0.991 | 48.563 |
| 4:2:0 average | 12 | 0.991 | 48.324 |
| 4:1:1 average | 12 | 0.987 | 46.940 |
| 4:1:0 average | 10 | 0.983 | 45.887 |
| 3:1:1 average | 10 | 0.976 | 41.788 |
| 4:4:2 - 4:2:2 avg. | 16.758 | 0.996 | 52.467 |
| 4:4:2 - 4:2:0 avg. | 13.513 | 0.994 | 50.380 |
| 4:2:2 - 4:1:0 avg. | 11.136 | 0.989 | 47.949 |
| 4:2:0 - 4:1:0 avg. | 10.381 | 0.987 | 47.085 |

## Bibliografija


[1] „YCbCr," [Elektronski]. dostopno na: https://en.wikipedia.org/wiki/YCbCr. [Poskus dostopa 29. april 2021].

[2] N. Campos, „RGB to YCbCr conversion," [Elektronski]. dostopno na: https://sistenix.com/rgb2ycbcr.html. [Poskus dostopa 29. april 2021].

[3] M. J. McNamara, „Chroma Subsampling Explained: Giving a Little to Save a Lot" [Elektronski]. dostopno na: https://www.projectorcentral.com/chroma-subsampling-explained.htm. [Poskus dostopa 29. april 2021].

[4] „Chroma subsampling," [Elektronski]. dostopno na: https://en.wikipedia.org/wiki/Chroma_subsampling. [Poskus dostopa 29. april 2021].

[5] K.-L. Chung, J.-S. Cheng in H.-B. Yang, „Effective Chroma Subsampling and Luma Modification for RGB Full-Color Images Using the Multiple Linear Regression Technique," *IEEE Access PP(99):1-1,* 2020.