

Metody postupného přenosu obrazu

The methods of progressive
image transmission

Jakub Dvořáček

Bakalářská práce
2005



Univerzita Tomáše Bati ve Zlíně
Fakulta Technologická

Vložit oficiální zadání diplomové/bakalářské práce

Prohlašuji, že jsem na celé diplomové práci pracoval samostatně a použitou literaturu jsem citoval.

Ve Zlíně, 24. 05. 2005

.....

jméno diplomanta

Abstrakt

Použití a přenos videa v digitálním tvaru vyžaduje použití rozsáhlých pamětí a velkých šířek přenosových pásem. Proto když chceme uchovat výhody digitálního tvaru informace a omezit náklady na uchování a přenos informací, musíme aplikovat metody komprese videa. Komprese umožňuje efektivní reprezentaci videosignálů použitím redukovaného počtu bitů.

Using and transmission of digital video requires using of large memories and large width of transmission bands. In case hold on digital information advantages and reduce of cost and transmission of information, it is possible to apply videocompression methods. Compression gives effective representation of video signal using reduced number of bits.

OBSAH

ÚVOD.....	6
1 ZPRACOVÁNÍ OBRAZOVÝCH SIGNÁLŮ.....	7
1.1 ZÁKLADNÍ SCHÉMA	7
1.2 ZÁKLADNÍ POJMY	8
1.3 TYPY OBRAZŮ	9
1.4 KÓDOVÁNÍ OBRAZOVÝCH SIGNÁLŮ	9
2. MOTIVACE VIDEOKOMPRESNÍ	10
3 PROBLÉMY VIDEOKOMPRESNÍ.....	13
3.1 PROBLÉM ÚZKÉHO ZORNÉHO POLE.....	14
3.2 PROBLÉM IDENTITY	15
4 VIDEOKOMPRESNÍ STANDARDY	17
4.1 MOTION JPEG	18
4.2 STANDARD MPEG	19
4.3 STANDARDY H.261 A H.263	23
5 ZÁKLADNÍ POJMY A DEFINICE.....	28
6 REPREZENTACE OBRAZU.....	30
6.1 HEURISTICKÉ REPREZENTACE OBRAZU	30
6.2 HIERARCHICKÉ REPREZENTACE OBRAZŮ	31
7 PYRAMIDY A JEJICH ROZDĚLENÍ	33
8 NĚKTERÉ KONKRÉTNÍ TYPY PYRAMID	35
8.1 NEPŘEKŘÝVAJÍCÍ SE PYRAMIDA 2×2	35
8.2 PŘEKŘÝVAJÍCÍ SE PYRAMIDA 4×4	36
8.3 GAUSSOVA PYRAMIDA	36
8.4 LAPLACEOVA PYRAMIDA	37
8.5 HRANOVÁ PYRAMIDA.....	37
8.6 PŘEKŘÝVAJÍCÍ SE PYRAMIDA 2×2, PLOCHA RASTRU S FAKTOREM 2	38
8.7 PYRAMIDY S HEXAGONÁLNÍM RASTREM.....	39
9 POROVNÁNÍ JEDNOTLIVÝCH TYPŮ PYRAMID.....	40
10 ALGORITMUS SEGMENTACE OBRAZU POMOCÍ PŘEKŘÝVAJÍCÍ SE PYRAMIDY	41
11 BUDOVÁNÍ PŘEKŘÝVAJÍCÍ SE PYRAMIDY	42
11.1 VARIACE PROCESU.....	44
11.1.1 INICIALIZACE.....	44
11.1.2 VÝBĚR OTCE.....	45
11.1.3 NEROZHODNOST	46
11.1.4 SEKVENCE ITERACÍ.....	46
11.1.5 VÁŽENÉ LINKOVÁNÍ	46
11.1.6 INTERAKCE UVNITŘ HLADIN.....	47
12 ZÁVĚR.....	49
SEZNAM POUŽITÝCH ZDROJŮ	50
SEZNAM OBRÁZKŮ.....	51
SEZNAM TABULEK	52

ÚVOD

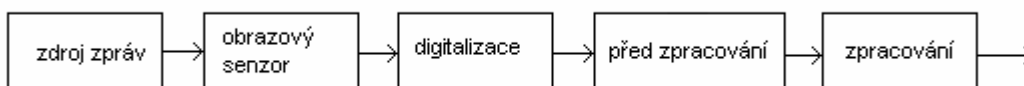
V průběhu posledních dvaceti let zpracování obrazových dat bylo předmětem intenzivního výzkumu a jako významná inženýrská aktivita při zavádění do praxe. Hlavními důvody je hromadné nasazení číslicových systémů a počítačů, zobrazovacích, přenosových zařízení, komerční elektroniky a pod. ve všech oblastech života společnosti. Tyto skutečnosti způsobili nevyhnutelnost číslicové reprezentace obrazové informace. Při současné a očekávané informační explozi je potřebné zpracovat velké množství informací, přičemž bude neustále kladený důraz na záznam, cenu a kvalitu těchto informací. Protože víc jak 70% informací přijímá člověk zrakem, je na místě připomenout přísloví “ Lepší jednou vidět jak desetkrát slyšet ”. Obrazy se vyznačují značným množstvím číslicových dat, proto by bylo velmi nákladné cenově a technicky je přenášet a zaznamenávat bez komprese. Oblasti aplikací komprese obrazových dat a nebo kódování obrazu jsou skutečně velmi široké. Nejvíc pozornosti bylo věnované oblasti videotelefonie, telekonferencím vzhledem k jejich komerčnímu využití pro poskytovatele telekomunikačních služeb. Další, v současnosti velmi zajímavá je oblast multimediálních aplikací, která zahrnuje např.: telekomunikace, rozhlasové a televizní vysílání HDTV, počítače, spotřební elektroniku. Vzhledem k širokému spektru takto používaných číslicových zařízení různých výrobců, kteří je poskytují, bylo nutné kvůli kompatibilitě vytvořit a přijat mezinárodní normy, doporučení a standardy pro kompresi různých typů obrazů v uvedených oblastech. V rámci ISO/CCITT už byli přijaté standardy JPEG (kódování a komprese statického obrazu), CCITT/H.261 (oblast videokonferencí) , MPEG (komprese a kódování pohyblivého obrazu) , MHEG (komprese a kódování obrazu pro multimediální aplikace) . Pyramidální reprezentace obrazu je v poslední době věnovaná pozornost v oblasti zpracování obrazové informace, robotiky, počítačové grafiky. Důvodem tohoto zájmu je skutečnost, že díky hierarchickému uspořádání je možné v úlohách obraz zpracovávat jednoduchými a přitom vysoce efektivními algoritmy.

1 ZPRACOVÁNÍ OBRAZOVÝCH SIGNÁLŮ

1.1 Základní schéma

Pod pojmem číslicové zpracování obrazové informace se rozumí převod obrazové zprávy do číslicového tvaru, zpracování v číslicovém tvaru a přenos resp. další zpracování obrazové informace.

Všeobecný model číslicového zpracování obrazové informace je možné znázornit blokovým schématem na obr.1.



Obr. 1. Všeobecný model číslicového zpracování obrazové informace.

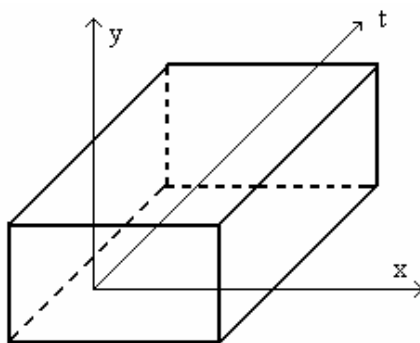
- Zdroj zpráv produkuje jednotlivé obrazové zprávy a je možné ho považovat za zdroj náhodného procesu, v kterém se jednotlivé zprávy (obrazy) vyskytují z různou pravděpodobností a statistickou závislostí.
- Obrazový senzor transformuje obrazovou zprávu nejčastěji pomocí fotoelektrické konverze na analogový elektrický signál.
- V bloku digitalizace se tento signál blokuje, kvantuje a číslicově kóduje. Výsledkem procesu digitalizace je digitální obraz, který je reprezentovaný maticí čísel, reprezentujících jas obrazu.
- Předzpracování obrazové informace spočívá v odstranění různých zkreslení, šumů, kterými byl obraz znehodnocený v procesu snímání a digitalizaci. Součástí předzpracování je někdy též úprava obrazových dat (obrazu), která zvyšuje efektivnost dalšího zpracování, např. kódování.
- Zpracování obrazové informace závisí od cílu číslicového zpracování.
Může to být :

- **archivace obrazu** – je efektivní uložení obrazů na paměťových médiích tak, aby se dosáhlo co možná největší efekt komprese dat.
- **analýza obrazu** – má za cíl výběr určitých charakteristických vlastností (příznaků) obrazu, které můžou sloužit např. při rozpoznávání obrazu.
- **přenos obrazů** – má za cíl efektivní a věrný přenos obrazů, které se na přijímací straně musí co nejméně lišit od původní obrazové zprávy.

Číslicové obrazové komunikační systémy dovolují komunikaci s obrazem (videm), přičemž v tomto systému prioritní informací je dynamický barevný obraz a jako přídatné informace jsou řeč, počítačové údaje, grafika a statický binární obraz.

1.2 Základní pojmy

Vizuální zpráva je možné si představit jako obrazový tok $B(x,y,t)$, kde funkce $B(x,y,t)$ vyjadřuje plošné rozdělení světelné veličiny probíhající v čase t , přičemž se pro názornost předpokládá obmezený geometrický rozměr obrazu v osách x,y (obr.2).



Obr. 2. Geometrický rozměr obrazu.

Když výchozí obraz je bez vnitřního pohybu, jde o tzv. statický obraz, který je možné popsat funkcí $B(x,y)$, t.j. obrazový tok je závislý jen od prostorových souřadnic a nemění se v závislosti od času. Naopak, funkce $B(x,y,t)$ popisuje dynamický (pohyblivý) obraz.

1.3 Typy obrazů

Binární obrazy - jas každého obrazového prvku je reprezentovaný jen jedním bitem („0“ a nebo „1“), čímž máme dvě jasové úrovně (černá – „0“, bílá – „1“).

Statické obrazy - je to obraz bez vnitřního pohybu (fotografie).

Jas každého obrazového prvku je reprezentovaný větším počtem bitů než jeden, přičemž získáváme větší počet jasových úrovní než dva.

Dynamické (pohyblivé)obrazy - představují obrazové sekvence - video. Je to postupnost statických snímků

1.4 Kódování obrazových signálů

Kódování je úprava vnitřní struktury zprávy z hlediska způsobu uložení informačního obsahu tak, aby tato struktura vyhovovala požadavkům informačního procesu. Obecně je kódování předpis, který každému prvku konečné množiny A přiřazuje jedno slovo v konečné množině B , přičemž množina A se nazývá zdrojová abeceda a její prvky jsou zdrojové znaky a množina B se nazývá kódová abeceda a její prvky jsou kódové znaky.

Cíle kódování :

1. Přizpůsobení formy uložení informačního obsahu z hlediska technického uspořádání přenosového kanálu (například číslíkové systémy potřebují binární vyjádření, člověk dekadické).
2. Zvýšení efektivity zprávy , t.j. odstranění nevyužitelné redundance.
3. Cílevědomé zvyšování užitečné redundance zprávy, které je zaměřené na zvýšení odolnosti vůči poruchám. Toto (tzv. kanálové) kódování zavádí záměrně do procesu kódování redundanci, která umožňuje odhalit a nikdy taky opravit chyby při přenosu (detekční a korekční kódy).

Kódy je možné rozdělit z více hledisek, přičemž rozeznáváme kódy :

1. **Rovnoměrné** (blokové) kódy, v kterých blokové znaky mají stejnou délku, t.j. stejný počet symbolů.
2. **Nerovnoměrné** kódy, ve kterých kódové znaky mají rozličnou délku.

2. MOTIVACE VIDEOKOMPRESSE

Pojem komprese je možné definovat ve smyslu efektivní reprezentace zdrojového signálu v digitálním tvaru. Tuto v podstatě stručnou definici můžeme rozšířit přidáním požadavků na zachování věrnosti daného signálu. Komprese je tedy metodou reprezentace zdrojového signálu v digitálním tvaru s nejmenším počtem bitů při zachování podmínky přijatelné ztráty věrnosti.

Zdrojovým signálem může být text, obraz, zvuk, video nebo jakýkoliv jiný signál, který potřebujeme přenášet a uchovávat. Metoda komprese může být bezztrátová (Huffmanovo, Shannonovo-Fanovo, aritmetické kódování) nebo ztrátová (predikční, transformační, subpásmové kódování). Při bezztrátové kompresi (lossless compression) je možné zrekonstruovat původní signál zdroje informací dokonale. Při ztrátové kompresi (lossy compression) se vychází obvykle z modelu vnímání informace uživatelem a podle požadavků na kvalitu, resp. rozsah vnímání se redukuje objem přenášené informace.

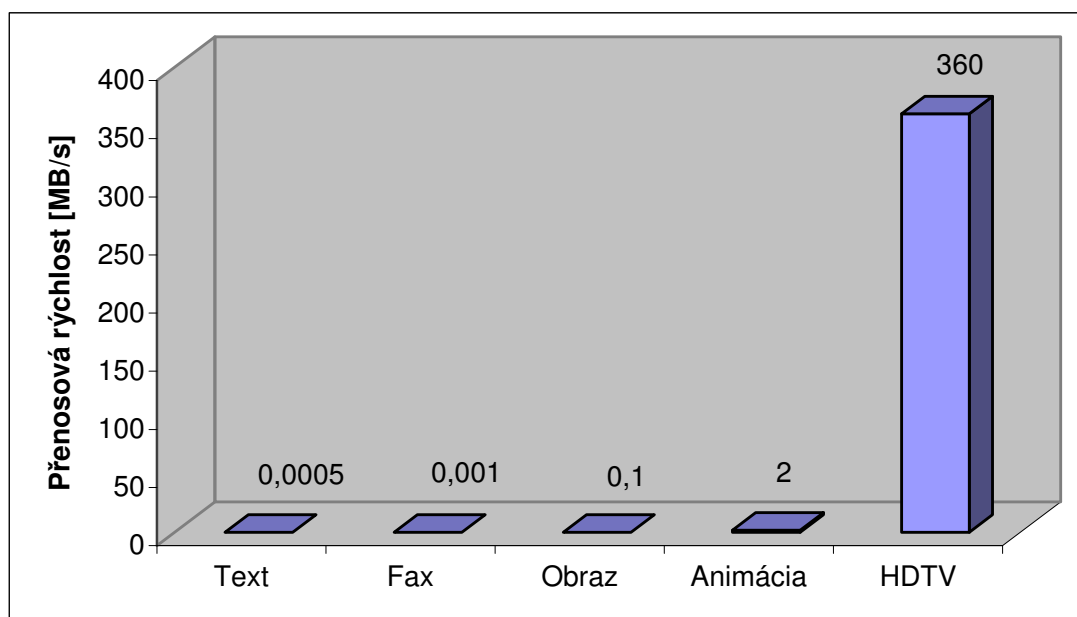
V tomto případě dokonalá zpětná rekonstrukce původního signálu není možná. Tyto metody ale umožňují dosáhnout vyšších kompresních poměrů, proto se často používají pro paměťově náročné média, jako je video a zvuk.

Pro kompresi údajů v současných informačních (obr. 3) a přenosových (obr. 4) systémech hovoří tři základní důvody :

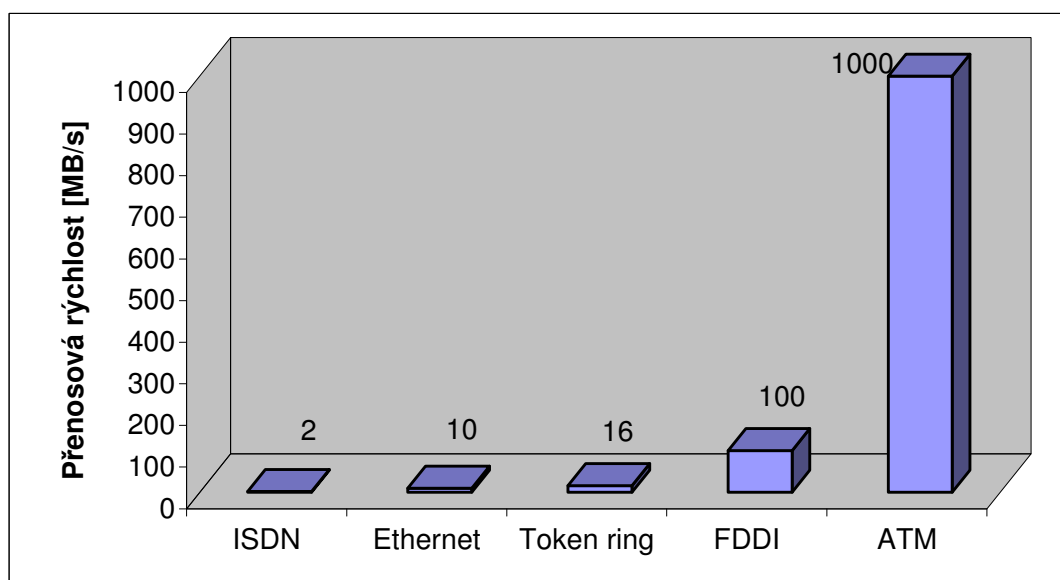
Značné paměťové nároky multimediálních dat, relativně pomalé paměťové zařízení, které neumožňuje přehrávání multimediálních dat (především videa) v reálném čase

Šířka pásma současných přenosových sítí, která neumožňuje přenos videa a často ani zvuku v reálném čase

Nedostatek potřebné šířky přenosového pásma, nedostatek velkokapacitních pamětí a médií na uchovávání informací (tab. 1) motivoval k nutnosti úspor použitím at' už jednodušších nebo složitějších technických postupů.



Obr. 3. Rychlost různých informačních technologií.



Obr. 4. Přenosová rychlost různých přenosových technologií.

Právě zmíněná šířka přenosového pásma a kapacita pamětí se staly neakceptovatelnými pro přenos video signálu. Jeden snímek s rozměry 640×480 obrazových prvků barevného videa s rozlišením 24 bitů na obrazový prvek má velikost přibližně 1 MB. Při snímkové frekvenci 30 snímků za sekundu a dvouhodinové velikosti filmu je objem informace až 215 GB!

I kdyby bylo k dispozici dostatečně velké paměťové médium, rychlost paměťových zařízení neumožňuje přehrávat video v reálném čase z důvodu nedostatečné rychlosti přenosu dat. Podle přecházejícího výpočtu by musela přenosová rychlost paměťového zařízení aspoň 20 MB/s. Avšak současné technologie umožňují přenosové rychlosti kolem 1.2 MB/s (CD-ROM technologie ~ 900 KB/s). Proto je při současném stavu těchto technologií zatím jedinou možností využití metod komprese před ukládáním dat na paměťové zařízení nebo přenosem dat a jejich následná dekomprese před přehráváním.

3 PROBLÉMY VIDEOKOMPRESSE

Při přenosu pohyblivého obrazu je jedním ze zdrojů redundance přenos informací o těch částech obrazu, které se nemění. Příčinou mezisnímkové difference je pohyb v obrazu, změna osvětlení, pohyb kamery nebo šum. Z uvedeného vyplývá, že dynamický obraz obsahuje oblasti s pohybem nebo bez pohybu. Mezisnímkové difference, chybové obrazové prvky v oblasti bez pohybu resp. s malým pohybem jsou zanedbatelné a nemusí se kódovat resp. vysílat.

Pokud vycházíme z předpokladu, že při pohybu objektů v pozorované scéně dochází ke změně úrovně jasu v místě pohybu, potom nejjednodušším způsobem detekce pohybu se jeví metoda rozdílového obrazu. Princip je možné zapsat podle následujícího vztahu :

$$d_{1,2}(i, j) = \begin{cases} 0 & \text{pro } |f_1(i, j) - f_2(i, j)| < t \\ 1 & \text{jinak} \end{cases} \quad (1)$$

kde d je binární rozdílový obraz, v kterém hodnoty „0“ představují navzájem si odpovídající místa snímků f_1 a f_2 , v kterých nedošlo k významné změně jasové funkce mezi okamžikem snímání obou obrazů, hodnoty „1“ představují místa obrazů, v kterém došlo k významné změně. Na posouzení významnosti změny je použita konstanta t , která představuje práh detekce a potlačí šumy a jiné různé nepřesnosti způsobené při snímání kamerou. Pokud je práh vysoký, může způsobit neschopnost detekce pomalých pohybů nebo pohyby malých objektů. Tuto metodu nemůžeme zařadit přímo mezi metody estimace pohybu, protože pohyb jen detektuje.

Metoda kumulativního rozdílového obrazu vychází z metody rozdílového obrazu, avšak poskytuje i informace o směru a časových vlastnostech detekovaného pohybu.

$$d_{akum}(i, j) = \sum_l^n a_l |f_1(i, j) - f_l(i, j)| \quad (2)$$

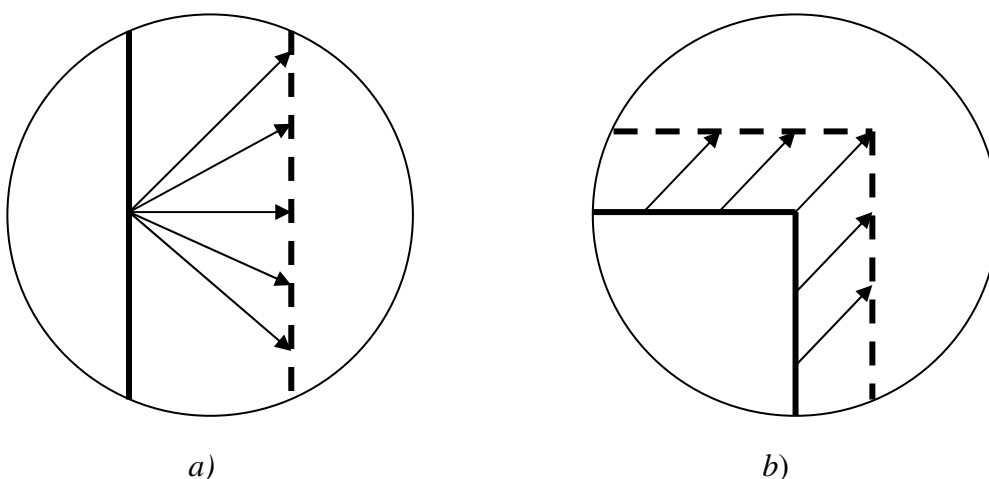
kde f_1 je referenční obraz, f_i je posloupnost snímků a_i jsou váhové koeficienty a n je počet snímků. Tyto metody však nemají všeobecnou platnost. Platí, že pohyb způsobuje časové změny jasové funkce v sekvenci snímků, ale neplatí to naopak t.j. že časové změny jasové funkce určují pohyb v obraze.

Při estimaci vektoru pohybu rozeznáme dva základní problémy s určením vektoru pohybu:

1. Problém úzkého zorného pole
2. Problém identity

3.1 Problém úzkého zorného pole

Na obr. 5a je znázorněna pohybující se hrana. V tomto případě neumíme jednoznačně určit vektor posunutí. Můžeme jen předpokládat, že má směr normální ke hraně. Tento problém je známý jako problém úzkého zorného pole, resp. problém apertury, který je způsobený lokálními operátory, časovými nebo prostorovými derivacemi, které pokrývají jen malý prostor nebo malou část objektu. Jednoznačně určit vektor pohybu je možné jen v případě, že se roh objektu nachází ve vnitřku pole pokrytí objektu (obr. 5b).

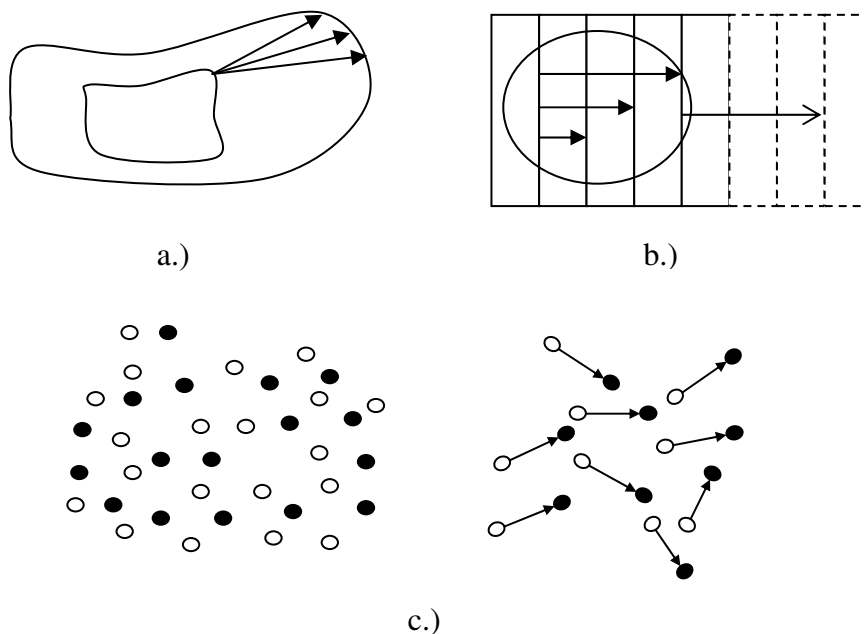


Obr. 5. Problém úzkého zorného pole resp. problém apertury.

3.2 Problém identity

Problém apertury vychází z toho, že není možné jednoznačně určit korespondující bod na hraně v následujícím snímku sekvence, protože není možné rozlišit body na hraně. Problém apertury je osobitým případem problému identity.

Všeobecně problém identity znamená, že není možné najít dva korespondující body v dvou po sobě následujících obrazech obrazové sekvence. Na obr. 6a je znázorněn dvojrozměrný objekt, který v průběhu sekvence roste. Je jasné, že není jednoznačně možné určit vektory posunutí na celé hraně objektu. Obr. 6b ukazuje nejednoznačnost při určování pohybu s periodickým vzorcem. Dalším příkladem problému identity je, že v obraze se nachází více objektů stejného typu. Pokud jsou vzdálenosti mezi objekty menší jako vektor posunutí, není možné jednoznačně vektor pohybu určit (obr. 6c). V tomto případě však můžeme zkrátit čas mezi jednotlivými snímky tak, aby bylo možné určit vektor pohybu jednoznačně.



Obr. 6. Problém identity.

Tyto příklady demonstrují základní problémy analýzy pohybu v obrazových sekvencích. Vyplynají hlavně z faktu, že obrazové sekvence nezobrazují realitu trojrozměrného světa jednoznačně pomocí dvojrozměrného rámce obrazu. To znamená, že reálná shoda nemusí být identická s vizuální shodou v obraze.

4 VIDEOKOMPRESNÍ STANDARDY

Asi od roku 1989 se video stalo součástí množství programových balíčků. Problémy s implementací videa jsou způsobované hlavně obrovskými nároky na šířku přenosového pásma, proto jak chceme přehrávat video v dostatečné kvalitě v reálném čase, stává se použití komprese videa nutností.

Pro digitální video bylo vyvinutých a implementovaných mnoho algoritmů komprese. Dosahované kompresní poměry se liší podle subjektivního měřidla dostatečně přijatelné kvality. Tab. 1 sumarizuje některé používané kompresní algoritmy společně s kompresními poměry udávanými v literatuře.

Tab. 1. Přehled algoritmů komprese videa.

<i>Algoritmus komprimace</i>	<i>Typický kompresní poměr</i>	<i>Charakteristika</i>
INTEL RTV indeo	3 : 1	Datový tok 128x240 je interpolovaný na 256x240. Barva je pod vzorkovaná v poměru 4 : 1. Využívá jednoduchý rozdíl snímků.
INTEL PLV	12 : 1	Používá datový tok 256x240 s vektorovou kvantizací a kompenzací pohybu. Komprese je možná jen se speciálním zařízením
IBM Photomotion	3 : 1	Je zvolená optimální 8 bitová paleta. Používá RLE kódování a rozdíl snímků.
Motion - JPEG	10 : 1	Používá 2-D DCT na zakódování jednotlivých snímků. Poskytuje dobré výsledky v reálném čase s laciným, ale specializovaným vybavením. Podporuje přímý přístup, protože nepoužívá rozdílové snímky.
Fraktály	10 : 1	Dobře aplikovatelné na kompresi přírodních scénérií, avšak vyžaduje velmi vysoký výpočtový výkon.
Wavelety	20 : 1	Používají se 2-D a 3-D wavelety. Komprese je vykonávána náraz pro celý snímek, čímž nedochází k blokovým efektům jako u algoritmů s DCT.
H.261 / H.263	50 : 1	Algoritmus aplikovatelný pro kompresi a dekompresi v reálném čase. Určený pro video telekomunikace. Založený na 2-D DCT a jednoduché estimaci pohybu mezi snímky.
MPEG	30 : 1	Používá 2-D DCT s estimací pohybu a interpolací mezi snímky. Složitý a náročný algoritmus při kompresi, ale použitelný při dekompresi v reálném čase.

Ideální metoda komprese videa by měla :

- konkurovat v dosáhnutých kompresních poměrech metodě MPEG,
- mít možnost přehrávat video bez nutnosti použití drahých technických zařízení,
- přizpůsobit se (postupným snižováním kvality) při přetažení přenosových sítí nebo při použití méně výkonného počítače,
- umožnit kompresi videa v reálném čase bez nutnosti použití nákladného zařízení.

Výkonné metody videokomprese umožnily dostupnými mnohé aplikace videa. Můžeme je rozdělit do následujících skupin:

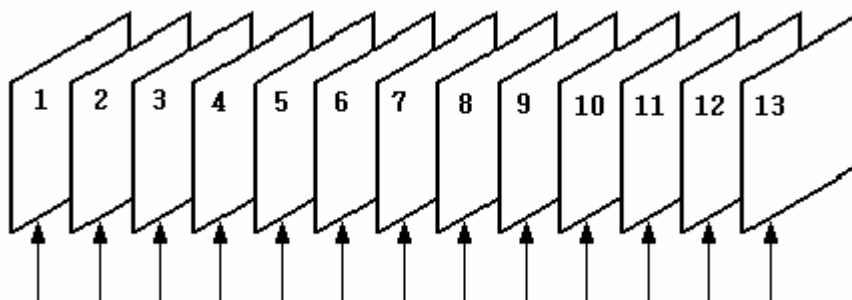
- a) *Televizní vysílání* zahrnuje dodávku digitálního videa do domácností. Vyžaduje jen malý počet kodérů a velký počet laciných dekodérů. Požadovaný kompresní poměr je $50:1$.
- b) *Aplikace přehrávání videa* jako jsou CD-ROM, Video CD, interaktivní hry si vyžadují též malý počet kodérů a velký počet laciných dekodérů. Odhadovaný kompresní poměr je asi $100:1$.
- c) *Počítačové videosystémy* zahrnují autorské systémy pro vytváření a editování videoprezentací, vyžadující stejně rychlé algoritmy kódování a dekódování. Odhadovaný kompresní poměr je v rozmezí $5:1$ až $50:1$.
- d) *Videokonferenční aplikace* vyžadují též symetričnost kódování a dekódování, přičemž hodnoty kompresního poměru jsou přibližně $100:1$.

4.1 Motion JPEG

Standard Motion JPEG je jednoduchým rozšířením standardu JPEG na dynamické obrazy. Motion JPEG zpracovává jednotlivé snímky ze sekvence individuálně použitím JPEG komprese. Motion JPEG je v podstatě velmi zjednodušená verze standardu MPEG, protože v něm absentuje estimace pohybu a mezi snímkové kódování. Je teda použité jen vnitro snímkové kódování.

I když kompresní poměry Motion JPEG nejsou tak vysoké jako u standardu MPEG, má Motion JPEG několik výhod. Hlavní nevýhoda MPEG-u je zdlouhavá a časově náročná estimace pohybu mezi snímky, která způsobuje, že kódovací proces MPEG-u je asi desetkrát delší jako u Motion JPEG-u.

Náhodný přístup do sekvence, rychlé přehrávání vpřed a vzad a jiné editovací operace v reálném čase vyžadují speciální požadavky, které jsou velmi dobře podporované v Motion JPEG-u, jako to ilustruje obr. 7.



Obr. 7. Přímý přístup do snímků sekvence M-JPEG.

Takovýto náhodný přístup není možný v každém snímku v sekvenci kódované pomocí MPEG-u. Navíc, když se kóduje každý snímek samostatně, nezávisle od ostatních, jsou nároky na paměť a výpočtový výkon kodéru Motion JPEG mnohem menší jako u kodéru MPEG. Nevýhodou Motion JPEG je to, že dosahuje kompresní poměr jen kolem 10:1 až 15:1.

4.2 Standard MPEG

Standard MPEG byl navrhnutý zejména na kompresi plně pohyblivého videa. Komprese využívá mezi rámcovou informaci a umožňuje dosahovat kompresní poměry až 200:1 tím, že uchovává jen změny mezi za sebou jdoucími snímky.

Standard MPEG-1 je zaměřený na kompresi videa s rozměry 320×240 obrazových prvků a přenosovou rychlostí 1 až 1,5 MB/s. MPEG-2 je určený pro vyšší kvalitu přenosu a je specifikovaný pro přenosové rychlosti od 2 do 80 MB/s. Podporuje vlastnosti standardů CCIR 601, EDTV a HDTV. MPEG-4 je určený pro kompresi plně pohyblivého videa složeného z malých rámců, které není třeba často obnovovat. Požadované přenosové rychlosti jsou 9 - 40 Kb/s a cílové aplikace jsou interaktivní multimédia a videotelefonie. Tento standard vyžaduje vývoj nových, na psychovizuálních modelech založených technikách kódování. Tab. 2 ilustruje různé formáty videa a odpovídající parametry MPEG. Algoritmus MPEG je možné použít pro aplikace vyžadující symetrii a asymetrii.

Tab. 2. Parametry algoritmu MPEG

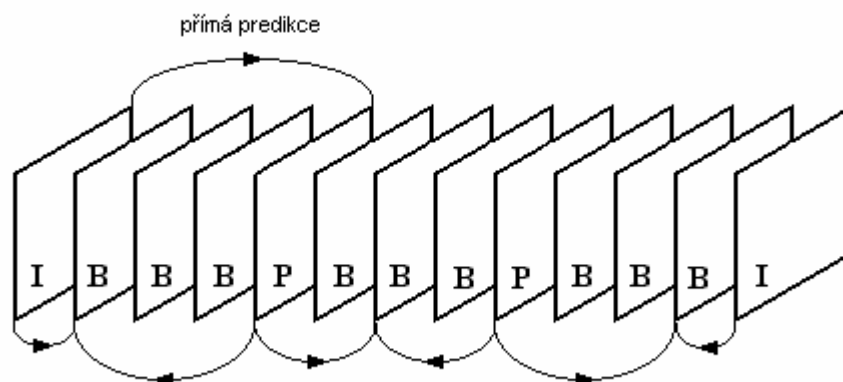
<i>Formát</i>	<i>Parametry videa</i>	<i>Datový tok po kompresi</i>
<i>SIF</i>	352x240 / 30 Hz	1,2 - 3 MB/s - MPEG-1
<i>CCIR 601</i>	720x486 / 30 Hz	5 - 10 MB/s - MPEG-2
<i>EDTV</i>	960x486 / 30 Hz	7 - 15 MB/s - MPEG-2
<i>HDTV</i>	1920x1080 / 30 Hz	20 - 40 MB/s - MPEG-2

Při návrhu standardu MPEG byly považované za nejdůležitější tyto vlastnosti:

- možnost přímého přístupu k obrázkům,
- možnost rychlého prohledávání vpřed a vzad,
- možnost špatného přehrávání,
- synchronizace zvuku a videa,
- robustnost vůči chybám,
- možnost editace,
- flexibilita formátů,
- cena.

Standard MPEG je složený ze třech částí, a to **systém** (synchronizace a multiplexní přenos videa a zvuku), **video** a **zvuk**.

Podle standardu MPEG jsou snímky kódované pomocí třech různých algoritmů, jak je to znázorněno na obrázku 8.



Obr. 8. Typy snímků standardu MPEG.

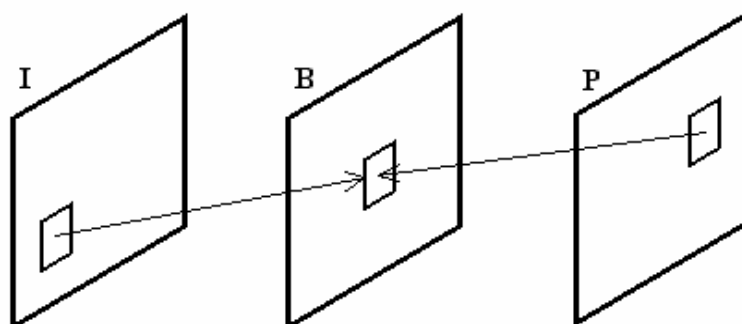
Snímky **I** (Intraframes) jsou samostatné, nezávisle zobrazitelné, bez potřeby informace z předcházejících nebo následujících snímků. V podstatě odpovídají formátu JPEG. Tyto snímky jsou kódované pomocí 2-D DCT pro bloky 8×8 obrazových prvků a kvantování koeficientů se provádí pomocí tabulky, která je přizpůsobena rozměrům (scaling). Tento mechanismus umožňuje použít adaptivní kvantování, které sice zvyšuje náročnost zpracování, avšak umožňuje zvýšit kompresní poměr až o 30 %. Kvantované koeficienty se zařadí do "cik-cak" postupnosti a kódují VLC kódem. Snímky I jsou použitelné pro přímý přístup do sekvence snímků a jejich kompresní poměr je nejnižší ze všech třech druhů snímků.

Snímky **P** (Predicted) jsou kódované pomocí dopředu predikčního kódování. Aktuální snímek je zakódován vzhledem na předcházející snímek typu I anebo P. Proces kódování je obdobný jako při standardě H.261, avšak s tím rozdílem, že přecházející vztažný rámec nemusí být fyzicky nejbližší předcházejícímu, jako je tomu při standardě H.261. Snímek P zavádí koncepci estimace a kompenzace pohybu mezi snímky. Každý makro blok je zakódovaný pomocí pohybového vektoru, který odhadne jeho předpokládanou polohu v předcházejícím snímku I anebo P. Kompresní poměr dosahovaný pro snímky typu P je výrazně vyšší, než u snímků I.

Snímky **B** (Bidirectional – obousměrné nebo interpolované snímky) jsou zakódované pomocí dvou referenčních rámců, předcházejícího a následujícího, které mohou být jen typu I nebo P. Na snímku B jsou jednotlivé makro bloky definované pomocí dvou vektorů. Zpětný vektor je referencí na makro blok v předcházejícím I nebo P snímku vektor je referencí na makro blok v následujícím snímku I anebo P. Z těchto důvodů jsou snímky I a P umístěné v zakódovaném toku dat před vlastními snímky B. Obousměrné (interpolační) kódování zajišťuje dosažení nejvyššího kompresního poměru.

Snímky I jsou vytvářené podobně jako při standardě JPEG, resp. M-JPEG. Snímky typu P a B jsou zakódované pomocí předcházejících a následujících snímků. Je odhadnutý vektor pohybu a je vypočítaná chyba mezi predikcí a aktuální hodnotou bloků. Predikční chyba je následně zakódovaná pomocí DCT společně s entropickým kódováním. Avšak v současném období je zkoumaný nástup waveletové transformace, při které nevzniká známý blokový efekt. Blokový diagram kodéru a dekodéru videa MPEG je na obr. 10.

Kódování snímku P a B zahrnuje použití estimace pohybu, kterou úlohou je najít nejvíc vyhovujících makro bloků v dostupných referenčních snímcích. Snímky P používají vždy jen dopřední predikci, zatím co snímky B využívají obousměrnou predikci, nazývanou též interpolací kompenzací pohybu, která je ilustrována na obr.9.



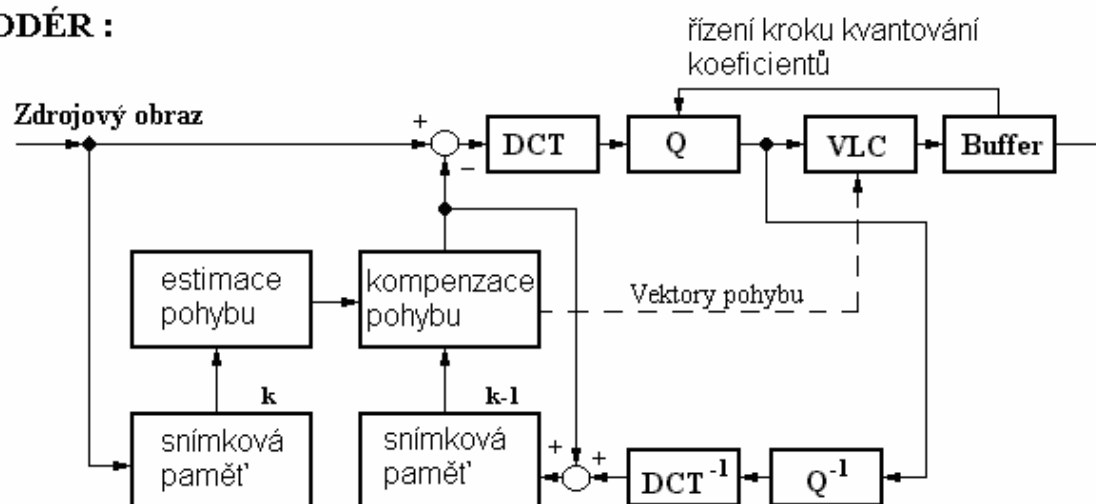
Obr. 9. Interpolace kompenzace pohybu implementovaná v standartě MPEG.

Pro snímky typu B může být použita predikce vpřed a zpět nebo interpolace. Makro blok v aktuálním B snímku může být odvozen z předcházejícího referenčního snímku ($B=A$ – dopřední predikce), z následujícího referenčního snímku ($B=C$ – zpětná predikce) nebo může být průměrem těchto bloků ($B=(A+C)/2$ - interpolace).

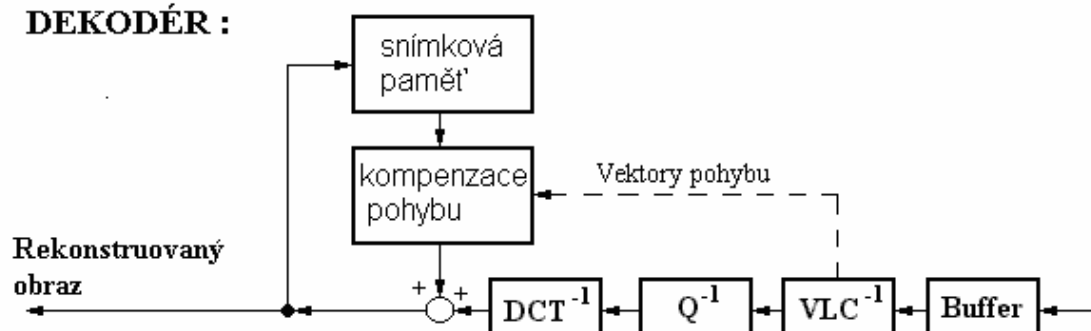
Pro každý makro blok snímku P a B je vypočítán jeden nebo dva vektory pohybu. Jeden vektor pohybu je potřebný pro snímek typu P a snímek typu B s do předu nebo zpětnou predikcí. Dva vektory pohybu jsou potřebné pro interpolované snímky B.

Standard MPEG nedefinuje přesně metodu estimace pohybu. Předpokládá však použití metod porovnávání bloků (block matching). V typických systémech MPEG je rozměr bloku 16×16 obrazových prvků (makro blok) a parametr p má hodnotu 6.

KODÉR :



DEKODÉR :



Obr. 10. Bloková struktura kódovacího a dekódovacího systému MPEG.

4.3 Standardy H.261 A H.263

ITU vyvinula video konferenční standard H.324 pro velmi nízké přenosové rychlosti. Doporučení H.324 zahrnuje 5 dokumentů:

- H.324 pro systém,
- H.223 pro multiplex,
- H.245 pro řízení,
- H.263 pro video,
- G.273 pro řeč.

H.261 umožňuje kódování videa pro přenosové rychlosti 64 Kb/s a víc, zatím co H.263, uvedený v doporučení H.324, umožňuje dosáhnout přenosovou rychlost kolem 16 Kb/s. Standard H.261/H.263, někdy jmenujeme i $p \times 64$ Kb/s, byl vytvořený za účelem dosáhnout velmi vysokých kompresních poměrů při přenosu plně barevného videa. Algoritmus $p \times 64$ je optimalizován na video telekomunikace tak, aby pokryl celou kapacitu kanálu ISDN ($p=1,2,\dots,30$). V důsledku limitované šířky pásma pro hodnoty $p=1$ a 2 se pomocí tohoto algoritmu dají implementovat jen aplikace typu videotelefon, t.j. zobrazení hlavy hovořícího (scéna typu head and shoulder). Pro p větší jako 6 je možné přenášet složitější obrázky a algoritmus se stává dobře použitelným pro videokonferenci. Protože tyto typy aplikací obvykle nejsou pohybově velmi intenzivní, používá standard H.261/H.263 jednoduchou estimaci pohybu v obrazu, čímž se dosahuje vysokých kompresních poměrů. Pro standardní obrázky je možné dosáhnout vysokých kompresních poměrů. Pro standardní obrázky je možné dosáhnout kompresní poměr 100:1 až 2000:1.

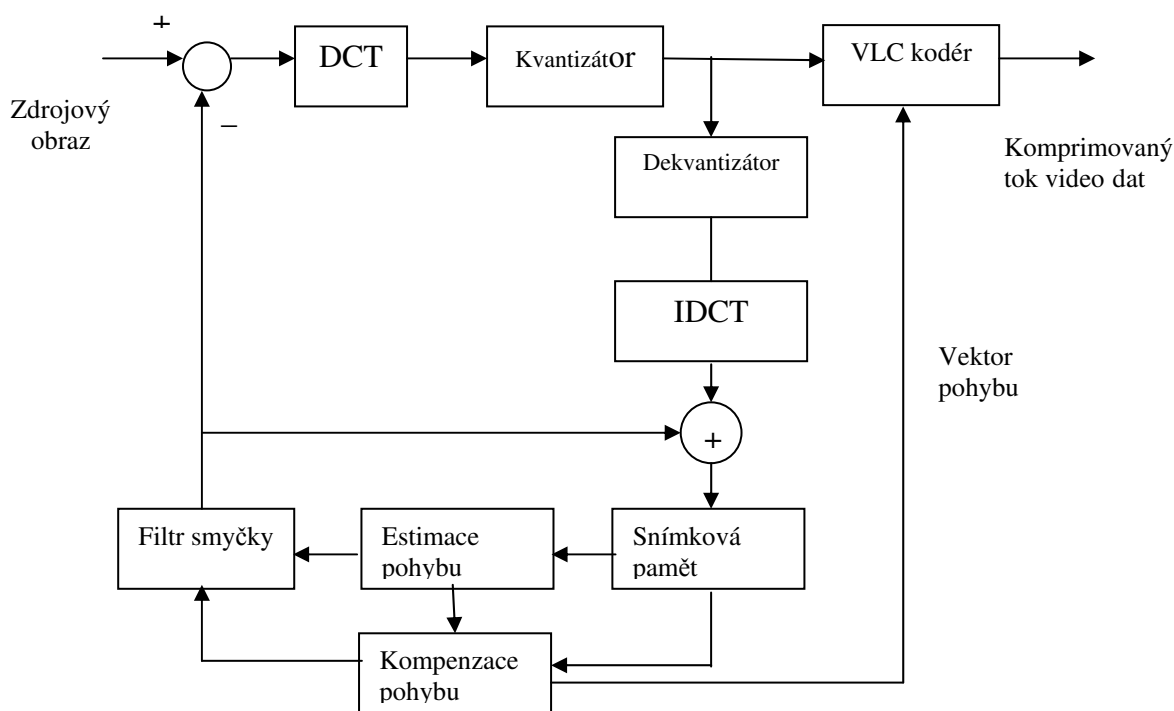
Pro algoritmus H.261 jsou definované dva druhy formátu obrazu: CIF (Common Intermediate Format) a QCIF (tab. 3), pro H.263 jsou specifikované i další obrazové formáty, a to SQCIF, 4CIF a 16CIF. Barevné obrazy jsou kódované pomocí jednoho jasového signálu a dvou rozdílových barvonosných signálů (model $YCbCr$) podle standardu CCIR 601. Složky C_b a C_r jsou vzorkované s faktorem 2 v horizontálním i vertikálním směru.

Tab. 3. Formáty obrazu pro video kodéry standardů H.261 a H.263.

Formát obrazu	Počet OP jasu	Frekvence snímků	Datový tok zdroje videa	Průměrný datový tok po kódování	kodér H.261	Kodér H.263
SQCIF	128 x 96	30 Hz	1,3 MB/s	26 Kb/s	Volitelný	Požadovaný
QCIF	176 x 144	30 Hz	9 MB/s	64 Kb/s ($p \times 64$ Kb/s)	Požadovaný	Požadovaný
CIF	352 x 288	30 Hz	36 MB/s	384 Kb/s ($p \times 64$ Kb/s)	Volitelný	Volitelný
4CIF	704 x 576	30 Hz	438 MB/s	3 - 6 MB/s	-	Volitelný
16CIF	1408 x 1152	30 Hz	2,9 GB/s	20 -60 MB/s	-	Volitelný

Algoritmus komprese videa kombinuje metody vnitrosnímkové a mezisnímkové komprese videa. Umožňuje rychlé zpracování v reálném čase. Algoritmus vytváří dva typy snímku:

- Snímek I, odpovídající vnitrosnímkové kompresi s použitím postupu podobného algoritmu JPEG, společně s využitím 2-D DCT, kvantizace a entropického kódování.
- Snímek P, získaný mezisnímkovým predikčním kódováním založeným na estimaci pohybu a rozdílové pulzní kódové modulaci (DPCM).



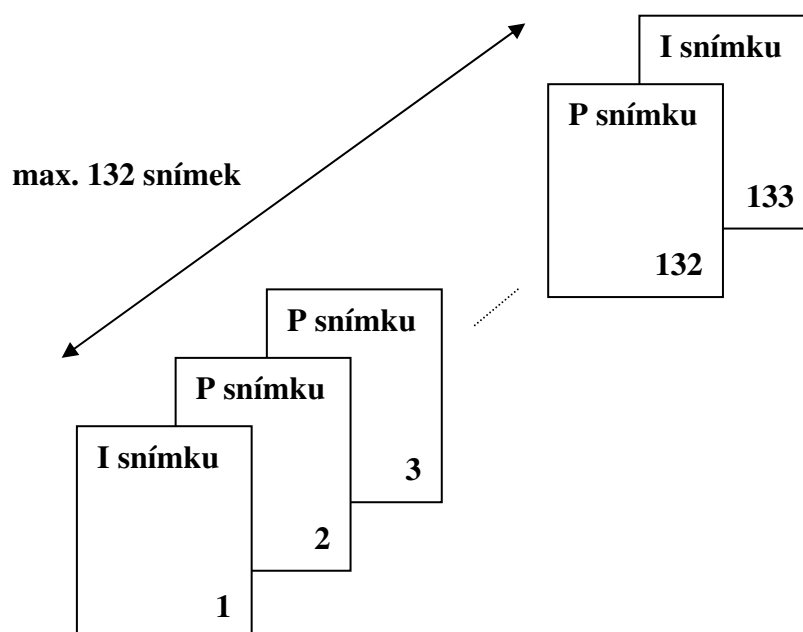
Obr. 11. Blokový diagram kodéru videa pro H.261/H.263.

Blokový diagram videokodéru je na obr. 11. Algoritmus kódování začíná zakódováním snímku I vnitro snímkovým algoritmem kódování, který se vyše do multiplexního kodéru videa a zároveň je dekódován v inverzním kvantizátoru a inverzní 2-D DCT. Dekódovaný snímek se uloží do snímkové paměti pro jeho využití při mezi snímkovém k kódování P snímku.

Při vnitro snímkovém predikčním kódování se porovnává každý makro blok aktuálního snímku s makro bloky z předcházejícím snímkem. Aby se zamezilo vzniku zpoždění při kódování, je k predikci použit vždy jen nejbližší předcházející rámeček. Rozdílový snímek se zakóduje pomocí 2-D DCT, je kvantován a vyslaný do multiplexního kodéru videa. Pro kódování snímku se použije jeden z uvedených postupů:

- Kódování DPCM bez kompenzace pohybu (nulové vektory pohybu),
- Kódování DPCM s nenulovými vektory pohybu,
- Bloky jsou filtrované volitelným předdefinovaným filtrem pro odstranění vysokofrekvenčního šumu.

V posledním kroku se vytvoření kompaktního kódu použije metoda kódování proměnlivou délkou kódu (VLC), např. Huffmanovo kódování. Vysoká komprese při standardech H.261 a H.263 se dosahuje i tím, že mezi dvěma I snímky může být maximálně 131 snímků P obr. 12.



Obr. 12. Sekvence snímku v standardech H.261/H.263.

Hlavní rozdíly mezi standardy H.261 a H.263 :

- H.263 využívá estimaci pohybu s přesností na $1/2$ obrazových prvků naproti celočíselné u H.261.
- H.261 podporuje jen 2 obrazové formáty QCIF a CIF, zatím co H.263 podporuje navíc ještě 3 obrazové formáty SQCIF, 4CIF a 16CIF.
- H.263 využívá mód "neomezeného" vektoru pohybu (Unrestricted Motion Vector Mode), kde se využívá zvětšená vyšetřovaná oblast a tedy jsou použité větší pohybové vektory. a zvýšená vyšetřovaná oblast a teda jsou použité vyšší pohybové vektory. Tento mód je speciálně používán v případě pohybu kamery.
- H.263 využívá zdokonalený predikční mód (Advanced Prediction Mode). Tento mód používá místo jednoho vektoru pohybu pro makro blok 16×16 op čtyři vektory pohybu pro submakro bloky 8×8 daného makro bloku. Čtyři vektory pohybu sice vyžadují více bitů, avšak dosažená predikční chyba je značně menší.
- Naproti standardu H.261, ve kterém se používá Huffmanově kódování, standard H.263 používá aritmetické kódování, které má asi o 10 % lepší účinnost.

5 ZÁKLADNÍ POJMY A DEFINICE

Operace, kterou si získá obrazový prvek na vyšší hladině, rozhoduje o vlastnostech a použití pyramidy. Příkladem takovéto operace je aritmetický průměr, respektive vážený aritmetický průměr

Překrývatelnost vyjadřuje způsob vytváření obrazového prvku na vyšší hladině ze skupiny obrazových prvků na nižší hladině. Podle překrývatelnosti rozlišujeme překrývající se a nepřekrývající se pyramidy.

Nepřekrývající se pyramida je taková, ve které skupina obrazových prvků na nižší hladině vytváří jen jediný obrazový prvek vyšší hladiny, teda každý syn má jen jednoho otce.

Překrývající se pyramida se vyznačuje překrýváním skupin obrazových skupin vytvářejících obrazové prvky na vyšší hladině, teda každý syn může mít více otců.

Velikost okna je velikost bloku obrazových prvků, podílejících se na tvorbě obrazového prvku vyšší hladiny.

Uzel každý obrazový prvek, nacházející se na libovolné hladině pyramidy.

Vrcholový uzel, uzel, který se nachází na nejvyšší hladině pyramidy.

Otec - obrazový prvek P – obrazový prvek nacházející se na bezprostředně vyšší hladině pyramidy a který překrývá obrazový prvek P. Otcovský uzel se označuje také jako uzel terminální.

Syn otce – obrazový prvek nacházející se na bezprostřední nižší hladině, který je překrytý otcem. Synovský uzel se označuje jako uzel neterminální.

Směrník (pointr)- charakteristika syna, která mu přiřazuje jednoho z potenciálních otců.

Linkování proces vytváření směrnic (pointrů).

6 REPREZENTACE OBRAZU

Vzhledem k cílům číslicového zpracování obrazu je kódování jen jednou z více operací v tomto zpracování. Proto se z tohoto aspektu zavedl pojem reprezentace obrazu, který má širší význam než kódování. Teda kódování, resp. efektivní kódování je speciálním případem reprezentace obrazu.

6.1 Heuristické reprezentace obrazu

Využívají statistické vlastnosti obrazu a využívají základní poznatky z teorie informací. Kódování vizuální informace si vyžaduje vzhledem na vysoký objem vstupních dat odstranit její nadbytečnou část, t.j. redundanci, která je v této informaci značně vysoká. Ve vizuální informaci je možné rozlišit dvě formy redundance :

- statistická redundance,
- subjektivní redundance.

Při heuristické reprezentaci obrazu se využívá blokové kódování binárních a víceúrovňových obrazů a využívá se také aproximační blokové kódování. Podle velikosti bloku a způsobu jejich dělení rozeznáváme :

1. Autoadaptivní jednorozměrné kódování.
2. Autoadaptivní dvojrozměrné kódování s jednorozměrným dělením.
3. Autoadaptivní dvojrozměrné kódování s dvojrozměrným dělením.

Při heuristických reprezentacích obrazu se využívá kódování délek úseků řádku. Využívají se :

Jednorozměrné RL kódy – se aplikují na jeden řádek obrazu, tedy kódují se jednotlivé úseky řádku bílé resp. černé barvy.

Dvojrozměrné RL kódy – se aplikují na víc řádků, tedy využívají taky vertikální korelaci mezi obrazovými prvky.

Jednorozměrné RL kódy je možné rozdělit do dvou skupin :

- lineární kódy,
- logaritmické kódy.

Do skupiny jednorozměrných RL kódů patří též tyto kódy :

- zkrácený Huffmanův kód,
- modifikovaný Huffmanův kód.

Dvojměrné RL kódy možné rozdělit do dvou základních skupin :

- současné kódování n řádků, se aplikuje na skupinu n řádků, přičemž počet řádků se volí v rozmezí 2 až 3,
- sekvenční kódování řádek po řádku, využívá obsah předcházejícího už zakódovaného řádku a aplikuje se postupně řádek po řádku.

Metody sekvenčního kódování řádek po řádku je možné rozdělit na :

- predikční kódování,
- kódování hraničních bodů.

6.2 Hierarchické reprezentace obrazů

Využívají topologické vlastnosti a hierarchickou strukturu obrazu. Zakládají se na principu rekurzivní dekompozice datového prostoru, přičemž je možné je rozdělit z těchto hledisek :

- dimenze datového prostoru,
- způsob dekompozice (regulární a neregulární, adaptivní a neadaptivní),
- adaptivnost dekompozice.

Při hierarchické reprezentaci obrazu se využívá :

1. **Kvadrantový strom obrazu (KSO)** - slouží na reprezentaci binárních obrazů, které mají tvar segmentu, přičemž pod pojmem segment se rozumí uzavřený obrazec, tvořený černými obrazovými prvky (úroveň „1“).

Kvadrantový strom obrazu je možné reprezentovat dvěma základními způsoby :

- směrníkovou (pointrovou) reprezentací,
- bezsměrníkovou (bezpointrovou) reprezentací.

Poznáme tyto reprezentace KSO :

- lineární kvadrantový strom,
- QC reprezentace,
- Reprezentace kvadrantového stromu obrazu DF výrazem (DF algoritmus, reprezentace víceúrovňových obrazů DF)
- reprezentace kvadrantového stromu obrazu pomocí BF výrazů,
- reprezentace kvadrantového stromu obrazu pomocí CD výrazů.

2. **Pyramidální reprezentace** - jako máme originální obraz v čtvercovém rastru $2^N \times 2^N$, pak pyramida je postupnost $\{ P_i \}$, přičemž P_{i+1} představuje obraz, kterého rozměr (rastr) je nějakou částí rozměru obrazu P_i .

Pyramidy je možné rozdělit podle těchto hledisek :

- typ operace,
- překrývatelnost (překrývající a nepřekrývající se pyramida),
- velikost okna.

7 PYRAMIDY A JEJICH ROZDĚLENÍ

Jestliže máme originální obraz (podle $P(n)$) zakódovaný v čtvercovém rastru $2^N \times 2^N$, potom pojmem pyramidu označujeme postupnost polí $\{ P_i \}$ takových, že P_{i-1} představuje obraz, kterého rozměry co do počtu obrazových prvků jsou nějakou částí rozměru obrazu P_i .

Nejnižší hladina pyramidu (základní) odpovídá originálnímu obrazu, každá další hladina je odvozena z předešlé určitou konstantní operací a to postupně až po nejvyšší hladinu.

Každý obrazový prvek vyšší hladiny je výsledkem určité aritmetické (logické) lokální operace nad ohraničenou skupinou obrazových prvků nižší hladiny. Příkladem takovéto operace je aritmetický průměr, či vážený aritmetický průměr.

Mimo už vzpomínaného typu operace, pomocí které se ze skupiny obrazových prvků nižší hladiny získá obrazový prvek vyšší hladiny, existují i další kriteria na odlišení typů pyramid.

Jedním z nich je překrývatelnost. Nepřekrývající pyramida je taková, v které vždy skupina obrazových prvků nižší hladiny vytváří jediný obrazový prvek vyšší hladiny, tedy každý syn má jednoho otce. Tento typ pyramidy je v případě nepřekrývající se pyramidy 2×2 vlastně úplným kvadrantovým stromem, kde rekurzivní dekompozice pokračuje až na úroveň obrazových prvků originálního obrazu. Překrývající se pyramida je na rozdíl od nepřekrývajících charakterizována překrýváním se jednotlivých bloků vytvářejících obrazových prvků na vyšší hladině, tedy každý obrazový prvek může mít několik otců. Dalším kriteriem určujícím typ pyramidy je velikost bloku, který se podílí na tvorbě zodpovídajícího obrazového prvku na vyšší hladině. Nejpoužívanějším představitelem tohoto typu je překrývající se pyramida 4×4 , která je současně reprezentantem „sudých“ pyramid používaných především při segmentaci. Z lichých pyramid se pravděpodobně nejvíc používají bloky 5×5 , které při vhodně volených váhových koeficientech aproximují Gaussovu křivku, proto se tyto pyramidy nazývají Gaussovými. Dále rozlišujeme dělení pyramid na ty, které zobrazují hranice (hrany) v oblasti v obrazovém poli a ty, které zobrazují vnitřek oblasti.

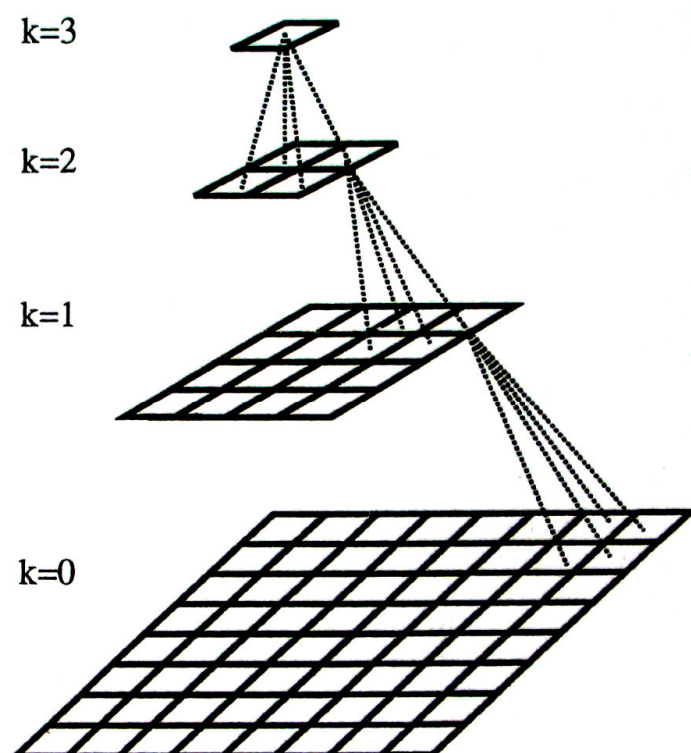
Pyramidy můžeme též rozlišit podle typu rastru, který je použitý na jednotlivých hladinách.

Nejčastěji používaný je klasický čtvercový a hexagonální, ostatní typy i když jsou principiálně možné, se používají zřídka. Mimo vzpomínaných základních typů existuje více zajímavých modifikací, které se používají kvůli některým svým výhodným vlastnostem.

8 NĚKTERÉ KONKRÉTNÍ TYPY PYRAMID

8.1 Nepřekrývající se pyramida 2×2

Klasická práce definuje pyramidu znázorněnou na obrázku 13. následujícím způsobem:



Obr. 13. Nepřekrývající se pyramida.

$$P_{k+1}(i,j) = 0,25 [P_k(2i, 2j) + P_k(2i, 2j+1) + P_k(2i+1, 2j) + P_k(2i+1, 2j+1)] \quad (3)$$

Což znamená, že každý obrazový prvek hladiny $k+1$ se získá jako jednoduchý aritmetický průměr hodnot od stínu šedé barvy jednotlivých obrazových prvků v odpovídajícím bloku 2×2 obrazových prvků na hladině k .

Jak má originální obraz rozměr $2^N \times 2^N$, potom nepřekrývající se pyramidu můžeme znázornit v podobě stromu, kterého kořenem je jediný obrazový prvek $P_N(I, I)$ a jeho listy jsou obrazovými prvky na hladině $k=0$.

V takto vytvořené pyramidě každý z obrazových prvků $P_K(I, I)$, s výjimkou obrazových prvků na vrcholu pyramidy má jednoho otce a každý $P_K(i, j)$ s výjimkou obrazových prvků originálního obrazu mají 4 syny $P_{K-1}(u, v)$, kteří se s výhrou označují podle světových stran – SZ, SV, JZ, JV. Jednotlivý synové potom podle tohoto označení mají souřadnice vypočítané ze souřadnic otce (i, j) na hladině k :

$$SZ = P_{k-1}(2i, 2j) \quad (4)$$

$$SV = P_{k-1}(2i, 2j + 1) \quad (5)$$

$$JZ = P_{k-1}(2i + 1, 2j) \quad (6)$$

$$JV = P_{k-1}(2i + 1, 2j + 1) \quad (7)$$

8.2 Překrývající se pyramida 4×4

Tento typ pyramidy je důkladně popsán v kapitole 11. (*Budování překrývající se pyramidy*) této práce.

8.3 Gaussova pyramida

Gaussova pyramida v podstatě představuje posloupnost obrazů (hladin), ve které každá další hladina je výsledkem aplikace dolně propouštějícího filtru na předcházející hladinu a filtrované frekvence jsou vždy posunuté o oktávu. Nejjednodušším způsobem dolně propouštějící filtrace je lokální zprůměrnování. V souladu s pravidly pro konstrukci číslicových filtrů, příspěvek obrazových prvků ležících blízko centra zprůměrnovaného okna je vyšší, než obrazové prvky ležící na jeho okraji, což je vyjádřené rozdílnými váhovými koeficienty v procesu zprůměrnování. Existuje velké množství rozličných operátorů (obvykle ve tvaru matice 3×3 , resp. 5×5), které jsou podrobené konvoluci se vstupním obrazem, výsledkem toho je matice představující vyfiltrovaný obraz.

8.4 Laplaceova pyramida

Je definována jako posloupnost pásmově filtrovaných obrazů, přičemž střední frekvence daného pásma je vždy o oktávu nižší než u předcházejícího. Každá hladina Laplaceové pyramidy vznikla jako rozdíl příslušných dvou hladin odpovídajících Gaussové pyramidě, jako rozdíl dvou dolně propouštějících filtrovaných obrazů. Protože operaci odpočítávání můžeme uskutečnit jen mezi obrazy (maticemi) stejných rozměrů, je nutné na vyšší hladiny Gaussové pyramidy aplikovat operaci EXPAND. Tato operace zdvojnásobí rozměry obrazu na příslušné hladině, a to tak, že mezi dva existující obrazové prvky umístí další, kterého hodnota je daná interpolací sousedních obrazových prvků.

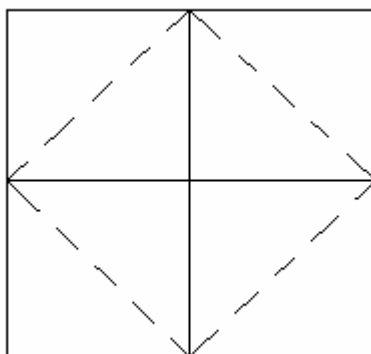
8.5 Hranová pyramida

Klasické pyramidy se dobře hodí na reprezentaci a analýzu plošných obrazů. Velmi často jsou však hlavními nositeli informace obrysy (hrany) objektů. Proto mají velký význam takové reprezentace, které dovolují výhody pyramid aplikovat na hrany a křivky. Nejjednodušším způsobem získání hranové pyramidy je aplikace hranového detektoru na všechny hladiny klasické pyramidy. Při tomto postupu však může nastat jev, že se hrana ztratí v důsledku zprůměrnování při přechodu na vyšší hladinu, případně, že se hrana vůči původnímu obrazu posune. Proto když chceme aplikovat takovýto postup musí po něm následovat linkování, při kterém se posuzuje míra věrnosti detekovaných hran dvou po sobě následujících hladinách. Jiný druh hranové pyramidy aplikuje hranový detektor ne na všechny hladiny klasické pyramidy, ale jen na originální obraz, přičemž na vyšší hladinu se přenáší jen ty hrany, které splňují určitá požadovaná kritéria.

8.6 Překrývající se pyramida 2×2, plocha rastru s faktorem 2

Jak je znázorněné na obr. 14. tato pyramida se vytváří tak, že se z každé čtveřice obrazových prvků na hladině k v pyramidě vytváří nová čtveřice na hladině $k+1$ a to tak, že hrany nové čtveřice na hladině $k+1$ jsou diagonálami obrazových prvků na hladině k . Každý obrazový prvek je tedy geometricky pootočený vůči svému synovi o 45° . Protože diagonály čtverce jsou $\sqrt{2}$ –násobkem délky hran, klesá rozměr každé další hladiny pyramidy směrem k vrcholu s faktorem $\sqrt{2}$ a její plocha s faktorem 2.

V takto vytvořené pyramidě je lehčí aplikovat některé lokální operace, které vyžadují pomalejší růst rozměru okna. Každý obrazový prvek má 2 otce a 4 syny, což dovoluje linkování a využívání vlastností překrývající se pyramidy.

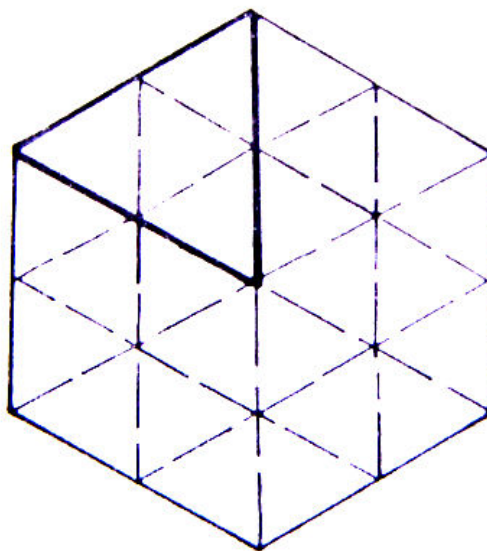


Obr. 14. Překrývající se pyramida.

8.7 Pyramidy s hexagonálním rastrem

Tato reprezentace se uplatní především tam, kde důležitější obrazové informace jsou lokalizované v centru obrazu, tedy např. v biologických vizuálních systémech. Hexagonální rastr disponuje touto výhodou, že všichni sousedé jsou stejně vzdáleni od centra a že například detekce hran se může uskutečňovat stejnoměrně v třech směrech. Hexagonální pyramidu můžeme vytvořit tak, že každý ze šestic trojúhelníků základního šestiúhelníku podrobíme dělení na čtyři menší, které reprezentují další hladiny pyramidy (obr. 15).

Mimo jiné klade hexagonální rastr nižší nároky na velikost paměti.



Obr. 15. Princip tvorby pyramid s hexagonálním rastrem.

9 POROVNÁNÍ JEDNOTLIVÝCH TYPŮ PYRAMID

Shrnutí vlastností jednotlivých typů pyramid můžeme vymezit oblast jejich použití. Nepřekrývající se pyramida 2×2 je svojí filozofií nejbližší kvadrantovým stromům, má malé nároky na velikost paměti. Má jednoduchou geometrickou strukturu, což se využívá při konstrukci pyramidálních procesů.

Překrývající se pyramidy mají sice o něco vyšší nároky na paměť a složitější geometrickou strukturu, avšak možnost podle potřeby měnit vazby mezi obrazovými prvky na různých hladinách poskytuje kvalitativní nové možnosti.

Gaussova a Laplaceova pyramida se používají tam, kde potřebujeme odděleně zpracovávat jednotlivé spektrální složky obrazu. Laplaceova je současně velmi úsporný prostředek kódování. Hranová pyramida aplikuje výhody pyramidální reprezentace na křivky a hrany, které jsou velmi často hlavním zdrojem informace.

Pyramida s hexagonálním rastrem je úspornější, detekci hran je možné uskutečňovat ve třech směrech a nejdělněji modeluje biologické vizuální systémy.

Z uvedeného je zřejmé, že každá reprezentace má specifické výhodné vlastnosti, avšak každá je dostatečně univerzální pro široké spektrum úloh.

10 ALGORITMUS SEGMENTACE OBRAZU POMOCÍ PŘEKRÝVAJÍCÍ SE PYRAMIDY

Cílem segmentace je abstrahovat z obrazu užitečnou informaci, či odlišit jednotlivé typy objektů navzájem i od pozadí. Protože se pracuje s reprezentací obrazů ve fyzikální rovině (na úrovni obrazových prvků) tj. s obrovským množstvím údajů, projeví se každé zlepšení algoritmů segmentace výrazným zvýšením rychlosti, resp. celé analýzy. Pyramidální reprezentace se už svojí filozofií, dovolující lokální operace obsáhnout libovolné velké okno obrazu, zařazuje mezi reprezentace mimořádně vhodné pro segmentaci. Když předpokládáme, že obrazové pole je tvořené několika druhy oblastí, se kterých každá má přibližně konstantní odstín (tón) šedé barvy, segmentaci takového obrazu můžeme vykonat prahováním úrovní šedé, tedy rozdělením stupnice šedé na intervaly a klasifikací každého obrazového prvku podle příslušnosti jeho šedého tónu k určitému intervalu šedé stupnice. Když je obraz zašuměný, takový proces segmentace vytváří mnoho chyb, protože šum způsobí, že některé obrazové prvky patřící k jednomu druhu oblastí mají tón, který leží v intervalu odpovídajícímu jinému druhu oblastí. Segmentace bude proto spolehlivější, když nejdříve obraz vyhladíme, abychom redukovali jeho šum. Obraz můžeme jednoduše vyhladit lokálním zprůměrováním. Tento proces však rozmáže hranice mezi oblastmi, když obrazové prvky v blízkosti hranice má sousedy ležící v té samé oblasti jako obrazové prvky a též sousedy v přilehlé oblasti. Jinými slovy kvalita procesů vyhlazení by se zlepšila, jak bychom mohli nejdříve obrazové pole segmentovat na oblasti tak, aby vyhlazení bylo vykonávané jen ve vnitřní určité oblasti, a ne přes její hranici. Jeden z přístupů k současné segmentaci vyhlazování umožňuje reprezentaci obrazů ve tvaru překrývajících se pyramid, která představuje posloupnost verzí obrazu s různým rozlišením, ve kterých jsou definované směrníky na navzájem se překrývajících prvky v posloupnosti jednotlivých hladin pyramidy. Vycházíme z toho, že v obraze s menším rozlišením (větším stromem), obrazové prvky ve vnitřku oblasti mají svoje tóny šedé méně zašuměné, protože tento obrazový prvek reprezentuje svým tónem průměr hodnot obrazových prvků v oblasti s vyšším rozlišením. Na druhé straně menší rozlišení způsobuje, že je méně pravděpodobné, že obrazové prvky leží ve vnitřku oblasti. Většina obrazových prvků překrývá dvě a nebo více oblastí. Reprezentace pomocí překrývajících se pyramid umožňuje využít současně výhody velkého a malého rozlišení.

11 Budování překrývající se pyramidy

Když má originální obrazové pole rozměr $2^N \times 2^N$ obrazových prvků, na to, abychom definovali verzi obrazu (hladinu) s menším rozlišením, použijeme exponenciální zužující pyramidu polí rozměru $2^{N-1} \times 2^{N-1}$, $2^{N-2} \times 2^{N-2}$..., tak, že k -tá hladina (verze) bude mít rozměr 2^{N-K} . Abychom předešli kolizím na hranicích jednotlivých polí, budeme je považovat za cyklicky uzavřené, tedy budeme uvažovat, že nalevo od prvního sloupce obrazových prvků na hladině s rozměrem 2^{N-K} bude umístěný poslední, tedy 2^{N-K-1} sloupec. Nad prvním řádkem bude řádek poslední.

Dále přiřadíme šedé tóny obrazových prvků na každé hladině ($k > 0$) vytvářením průměru šedých tónů bloků obrazových prvků velikostí 4×4 umístěných na hladině, která je pod ní. Bloky odpovídající přilehlým obrazovým prvkům se překrývají padesátiprocentně, proto tedy plocha hladin se stoupajícím k se redukuje s faktorem 4 a ne 16.

Jak chceme určit polohu obrazového prvku v pyramidě, musíme určit nejdříve hladinu, na které se nachází (index k) a polohu v rámci dané hladiny (indexy i, j). Když vezmeme například obrazový prvek (i, j, k) na hladině $k > 0$, bude překrývat blok prvků 4×4 s indexy

$$[u, v] [u, v+1] [u, v+2] [u, v+3] \quad (8)$$

$$[u+1, v] [u+1, v+1] [u+1, v+2] [u+1, v+3] \quad (9)$$

$$[u+2, v] [u+2, v+1] [u+2, v+2] [u+2, v+3] \quad (10)$$

$$[u+3, v] [u+3, v+1] [u+3, v+2] [u+3, v+3] \quad (11)$$

na hladině $k-1$, přičemž $u=2i-2$, $v=2j-2$, potom prvek $(i, j+1, k)$ bude překrývat blok obrazových prvků

$$[u, v+2] [u, v+3] [u, v+4] [u, v+5] \quad (12)$$

$$[u+1, v+2] [u+1, v+3] [u+1, v+4] [u+1, v+5] \quad (13)$$

$$[u+2, v+2] [u+2, v+3] [u+2, v+4] [u+2, v+5] \quad (14)$$

$$[u+3, v+2] [u+3, v+3] [u+3, v+4] [u+3, v+5] \quad (15)$$

na hladině $k-1$. Je zřejmé, že každý obrazový prvek pod hladinou vrcholu ($k < N-1$) je překrytý čtyřmi obrazovými prvky na vyšší hladině. Dále budeme prvky na hladině $k-1$,

kteřé odpovídají obrazovým prvkům na hladině k nazývat jeho syny a prvky na hladině k , které překrývají obrazové prvky na hladině $k-1$ otci. Z předcházejících úvah vyplývá, že každý obrazový prvek na hladině větší než nula má 16 synů, každý obrazový prvek na hladině menší než $N-1$ má 4 otce. Když na hladině $N-2$ je jen 16 obrazových prvků, každý z nich je synem všech čtyřech obrazových prvků na vrcholu pyramidy ($k=N-1$). Proces linování obrazových prvků je následující: obrazy (hladiny) s menším rozlišením jsou na začátku definované pomocí neváženého zprůměrnění tak, že hodnota otce je určena jako aritmetický průměr hodnot šedého tónu jeho synů. Vlastní linkování probíhá tak, že hodnota každého syna se porovná s hodnotami jeho čtyř otců a vytvoří se směrnik (resp. ukazatel, pointer), který ukazuje na nejpodobnějšího otce, toho, kterého hodnota je nejbližší hodnotě syna. Toto linkování se vykoná pro každý obrazový prvek a na každé hladině. Následuje přepočítání hodnot otců, přičemž hodnota otce se teď určí tak, že se zprůměrní hodnoty jen těch synů, kteří jsou k tomu-kterému otcovi přilinkovány. Jak je některý z otců „bezdětný“, můžeme mu přiřadit libovolnou hodnotu např. nulovou.

Na základě těchto nových hodnot otců vykonáme opět linkování, protože hodnoty otců se změnili a tím se mohli změnit i nejpodobnější otcové. Následuje opět zprůměrnění podle směrniků, potom linkování atd....

Bylo dokázané, že tento iterační proces konverguje a ustálí se až po několika (asi deseti) iteracích. Abychom vysvětlili tento proces, definujme bázi uzlu jako množinu prvků na nejnižší hladině (originální obraz), které jsou přilinkované (přes několik mezihladin) k danému uzlu. Na začátku procesu je základna každého uzlu čtvercovým blokem obrazových prvků. Když základna leží většinou uvnitř oblasti, je nejvíc pravděpodobné, že obrazový prvek se přilinkuje k uzlům na vyšší hladině, které leží také v této oblasti. Malé počáteční změny v průměrných hodnotách na vyšších hladinách směřují k tomu, že uzly nabývají hodnoty, které se blíží k průměrnému tónu šedé v této oblasti a nebo k více podobným oblastem (co do tónu šedé barvy) originálního obrazu. Předpokládejme, že na obraze nejsou více jak 4 druhy oblastí. Potom pro každý druh by měl být aspoň jeden obrazový prvek na vyšší hladině 2×2 pyramidy, kterého hodnota konverguje k průměru šedého tónu v oblasti daného druhu. Tento uzel bude přilinkovaný k obrazovým prvkům přes další obrazový prvek až obrazovým prvkům originálního obrazu, které leží ve vnitřku oblasti daného druhu. Jak víme, kolik druhů oblastí můžeme očekávat na obraze, můžeme potlačit některý z prvků na vrcholu pyramidy (zabránit, aby se k tomuto prvků přilinkoval nějaký syn) a ponechat jen tolik vrcholových prvků, kolik

druhů oblastí očekáváme. Případně můžeme spojit některé z těchto uzlů zprůměrněním jejich hodnot. Takovýmto způsobem můžeme zabezpečit, že počet stromů je stejný jako počet druhů oblastí originálního obrazu. Metoda selhává, když na obraze máme víc než 4 druhy oblastí. V tomto případě by prvky patřící dvou a nebo více druhů oblastí byli nucené přilinkovat se k jinému stromu.

Linkovací a zprůměrnující proces můžeme shrnout následovně:

- a) počáteční jednouché zprůměrnování šestnácti synů,
- b) linkování každého syna k jednomu ze čtyř otců, kterého hodnota je nejbližší k hodnotě syna,
- c) přepočítávání hodnot zprůměrnováním hodnot jen těch synů, kteří jsou k otcovi přilinkováni,
- d) změna směrnic podle nových hodnot,
- e) opakování kroků c a d dokud se hodnoty obrazových prvků ustálí.

V každém stavu tohoto procesu při linkování definuje množinu stromů (nejvíce 4) zakořeněných na vrchní hladině, odkud potom odvodíme hodnotu všech synů, kteří jsou přilinkováni k danému vrcholovému uzlu. Tento proces v konečném důsledku vyhladí obraz tím, že přiřadí každému obrazovému prvku jeho stromový průměr, tedy hodnotu vrcholového uzlu, který je kořenem stromu.

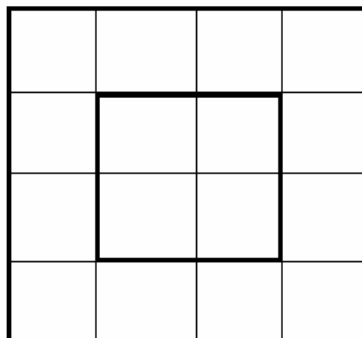
11.1 Variace procesu

Způsob přístupu k jednotlivým krokům procesu může být různý a každý z nich působí ve větší nebo menší míře na konečný výsledek. Možné variace se týkají toho, jak inicializovat hodnoty uzlů, jak vybrat nejpodobnějšího otce, jak postupovat v případě nerozhodnosti při linkování a jak postupovat při iteraci.

11.1.1 Inicializace

Základní metoda spočívá v počátečním přesměrování hodnot všech šestnácti synů. Možná obměna tohoto procesu spočívá ve výpočtu průměru jen ze čtyřech synů, kteří jsou otcovi nejbližší co do polohy, tedy leží v oblasti 2×2 v středu bloku 4×4 obr.16. Z ověření jednotlivých metod na různých druzích obrazu vyplývá, že při výpočtu průměrů ze čtyřech synů proces rychleji konverguje, dává lepší výsledky a klade menší nároky na paměť

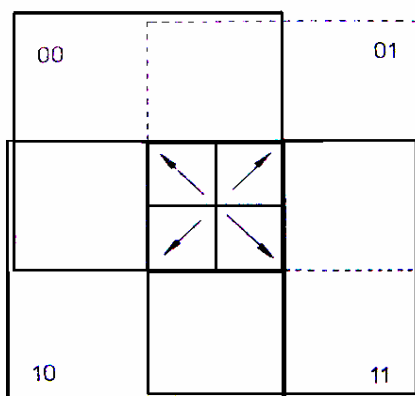
počítače, který obraz zpracovává. Je zřejmé, že když inicializaci vypočítáváme ze 16ti synů, tak při rozlišení 2^N úrovní odstínů šedé barvy, potřebujeme nejméně $N+4$ bity pro informaci o její hodnotě, v případě 4 synů potřebujeme $N+2$ bity, tedy o 2 bity méně.



Obr. 16. Výběr oblasti pro inicializaci.

11.1.2 Výběr otce

Každý uzel je v konečném důsledku přilinkovaný k otci, který má nejpodobnější hodnotu. Nerozhodnost vzniká, když všichni čtyři potencionální otcové (případně alespoň dva) mají stejnou barvu. Všimneme si, že pro každý uzel existuje jen jeden otec, jehož pozice vůči uzlu je neblíže, tedy uzel leží v centrální oblasti 2×2 . U všech ostatních otců leží v oblasti 4×4 (obr. 17, šipky synů ukazují na nejbližší otce). Tedy nerozhodnost v takovéto situaci odstraníme tím, že když máme více otců nejpodobnějších co do hodnoty, směrník volíme tak, aby ukazoval na nejbližšího otce co do polohy vůči synovi.



Obr. 17. Výběr otce.

11.1.3 Nerozhodnost

Jestliže dva otcové mají úplně stejné vlastnosti při linkování (barva nejbližšího otce je odlišná), máme možnost volby pořadí výběru otce např. *SZ*, *SV*, *JV* a *JZ*. Pořadí výběru neovlivňuje významné výsledky zpracování. Ve většině obrazů dává nejlepší výsledek náhodný výběr.

11.1.4 Sekvence iterací

V předešlém jsme iterovali tak, že jsme vypočítali průměrné hodnoty pro celou pyramidu, poté jsme celou pyramidu přilinkovali a počítali opět průměry podle linek pro všechny úrovně v pyramidě. Obměna tohoto procesu spočívá v tom, že budeme iteraci vykonávat jen mezi dvěma hladinami. Tedy vypočítáme průměry první hladiny, přilinkujeme syny originálního obrazu (nultá hladina) a ihned vypočítáme průměry pro otce z první hladiny podle směrnic a proces opakujeme. Po několika iteracích se hodnoty ustálí a přejdeme na výpočet mezi první a druhou hladinou atd., dokud nepřejdeme celou pyramidu.

11.1.5 Vážené linkování

Konstrukce linkované pyramidy, která byla dosud popsána, se zakládá na nuceném výběru, tedy každý syn byl nucený přilinkovat se k jednomu z potencionálních otců na vyšší hladině, k tomu, který byl nejpodobnější nebo nejbližší. Další možnou obměnou je použití váženého linkování, kde váhy spojené mezi uzlem a jeho potencionálními otci závisí na stupni podobnosti. V tomto případě, podobně jako při základní metodě, hodnota uzlů $h(i, j, k)$ se na začátku vypočítá neváženým zprůměrováním hodnot jeho synů:

$$h(i, j, k) = 1/16 \sum h(u, v, k-1) \quad (16)$$

kde u, v jsou souřadnice syna na hladině $k-1$.

Takto určíme hodnotu každého uzlu v pyramidě až po hladinu 2×2 . V dalším je každému ze čtyřech otců přiřazena váha uzlu, která závisí na jeho podobnosti k danému uzlu. Když pro $p=1,2,3,4$ je $d(p)$ absolutní hodnota rozdílu mezi hodnotami uzlů a a jeho p -tého otce, potom váha spojení $w(p)$ mezi uzlem a tímto otcem je definována vztahem:

$$w(p) = \frac{1/d(p)^2}{\sum_{q=1}^4 1/d(q)^2} \quad (17)$$

Takto definované váhy jsou nezáporné a jejich součet je roven 1. Jestliže je některá hodnota $d(p)$ nulová, potom přiřadíme $w(p)=1$ a $w(q)=0$ pro $q \neq p$.

Po vypočítání vah, hodnoty uzlů určíme jako vážené průměry:

$$h(i, j, k) = \frac{\sum w(i, j, u, v) \cdot h(u, v, k-1)}{\sum w(i, j, u, v)} \quad (18)$$

kde $w(i, j, u, v)$ je váha mezi uzlem a jeho potencionálním otcem (i, j, k) . Dále se opět vypočítají difference a nové váhy, poté se určí nové hodnoty uzlů. Proces opakujeme, dokud se hodnoty uzlů neustálí. Když na vrcholu pyramidy se nachází čtyři uzly, v pyramidě se vytvoří čtyři stromy spojením uzlů s nejvyššími váhami, co vede k segmentaci originálního obrazu na čtyři oblasti. Jestliže známe potřebný počet oblastí (≤ 4), můžeme vyloučit některé z vrcholových uzlů (nastavením jejich hodnot na takové, které se velmi liší od šedých tónů oblastí základního obrazu, tedy váhy spojení budou slabé), což bude vést k segmentaci na menší počet oblastí.

11.1.6 Interakce uvnitř hladin

V procesu linkování všechny interakce mezi uzly jsou vertikální. V tomto procesu přímo nezohledňujeme interakce mezi uzly na stejné hladině.

Důvodem pro využití interakcí mezi uzly na jedné hladině je fakt, že především na nižších hladinách pyramidy sousední uzly s velkou pravděpodobností patří do stejné části obrazu. Mezi nejzákladnější metody vyhlazování založené na vyrovnávání hodnot sousedních obrazových prvků.

Které patří k stejné oblasti, patří metoda mediální filtrace a metoda selektivního zprůměrnění. Při mediální filtraci je hodnota uzlu nahrazena mediánem hodnot sousedních uzlů. Jestliže uzel P leží ve vnitřku oblasti, medián je blízký k průměru, zatímco v případě, že uzel P leží v blízkosti hranice oblasti, medián má hodnotu, která se blíží hodnotě vypočítané jen z uzlů ležících ve stejné oblasti jako uzel P .

Metoda selektivního zprůměrnění vytváří průměr uzlu z hodnot r sousedních uzlů, kterých hodnota je nejbližší hodnotě uzlu P . Vybrané sousední uzly patří s nejvyšší pravděpodobností do stejné oblasti jako uzel P , dokonce i v případě, že se P nachází v blízkosti hranice. Tato metoda umožňuje dosáhnout výrazné vyhlazení bez rozmazání hranic.

Využití interakcí uvnitř hladin zlepšuje výsledky segmentace a dokonce umožňuje dosáhnout dobré výsledky i v případě použití inverzní lineární váhové funkce když se interakce uvnitř hladin využijí i při vytváření vah spojení a při vypočítávání hodnot uzlů. V tomto případě při vytváření váhy spojení mezi uzlem P a jedním z jeho potencionálních otců používáme jen ty sousedy uzlu P , kteří mají stejného potencionálního otce.

12 ZÁVĚR

Tato písemná práce obsahuje přehled základních metod estimace pohybu, které umožňují snížit přenosovou rychlost dynamických obrazů. Práce zhodnocuje současný stav v této problematice, přičemž uvádí základní myšlenky a problémy estimace pohybu obrazu spolu s jejich využitím v průmyslu komerční bezpečnosti a v dalších oblastech lidské činnosti.

Tyto metody se opírají o matematické modely, které popisují jasové změny v obrazu. Jejich nevýhodou je složitá technická realizace, protože vyžadují množství matematických operací včetně operací násobení. Za určitých podmínek je možné počet matematických operací snížit, avšak při větších posunutích klesá přesnost a správnost výpočtu.

Metody postupného přenosu obrazů popisují algoritmy technické realizace zpracování digitálních (obrazových) dat do tvaru vhodného pro jejich přenos. Postupný přenos obrazů z hlediska uživatele se projevuje tím, že v první iteraci se zobrazí hrubá aproximace obrazu, která se postupně vylepšuje. Tento přístup lze pozorovat i v některých Internetových aplikacích. Jiné uplatnění této skupiny metod představují obrazové databáze, kde je zapotřebí najít hledaný obraz, přičemž rozhodnutí, jestli je hledaný obraz požadovaný (ten, který jsme hledali), zůstává na uživateli. Metody postupného přenosu obrazu pak umožňují posouzení uživatele na základě zvoleného rozlišení, čímž se zkracuje doba přístupu k jednotlivým snímkům. Tyto metody lze implementovat i do průmyslu komerční bezpečnosti, kde je někdy zapotřebí prohledávat obrazové databáze a zjistit nejen (bezpečnostní) incident, ale i stav před ním, i po něm.

Pyramidální reprezentace představují velmi efektivní prostředek na zpracování obrazu, který umožňuje dosáhnout uspokojivé výsledky v širokém spektru úloh. Protože oblast jejich použití není omezena jen na zpracování obrazu, vznikly speciální architektury výpočetních systémů, které umožňují paralelní a tím i velmi rychlé zpracování. Jejich uplatnění v praxi má velký ekonomický a strategický význam.

SEZNAM POUŽITÝCH ZDROJŮ

ČAPEK, J., FABIAN, P.: *Komprimace dat - principy a praxe*. Praha: Computer Press, 2000.

KŘEČEK, S. a kol. *Příručka zabezpečovací techniky*. Blatná: Blatenská tiskárna, 2003.

MORKES, D.: *Komprimační a archivační programy*. Praha: Computer Press, 1998.

REPČÍK, D. ŠTULRAJTER, J. : *Číslicové filtry v subpásmovém kódování*. Liptovský Mikuláš: VVTŠ, 1996.

ŠTULRAJTER, J. , LEHOTSKÝ M. , CHMÚRNÝ J. *Subpásmové kódování obrazov pomocí banky filtrů*. Liptovský Mikuláš: VVTŠ, 1996.

SEZNAM OBRÁZKŮ

Obr. 1. Všeobecný model číslicového zpracování obrazové informace.....	7
Obr. 2. Geometrický rozměr obrazu.	8
Obr. 3. Rychlost různých informačních technologií.....	11
Obr. 4. Přenosová rychlost různých přenosových technologií.	11
Obr. 5. Problém úzkého zorného pole resp. problém apertury.	14
Obr. 6. Problém identity.	15
Obr. 7. Přímý přístup do snímků sekvence M-JPEG.	19
Obr. 8. Typy snímků standardu MPEG.	20
Obr. 9. Interpolace kompenzace pohybu implementovaná v standartě MPEG.	22
Obr. 10. Blokovaná struktura kódovacího a dekódovacího systému MPEG.	23
Obr. 11. Blokový diagram kodéru videa pro H.261/H.263.	25
Obr. 12. Sekvence snímku v standardech H.261/H.263.	26
Obr. 13. Nepřekrývající se pyramida.....	35
Obr. 14. Překrývající se pyramida.	38
Obr. 15. Princip tvorby pyramidy s hexagonálním rastrem.....	39
Obr. 16. Výběr oblasti pro inicializaci.....	45
Obr. 17. Výběr otce.....	45

SEZNAM TABULEK

Tab. 1. Přehled algoritmů komprese videa.	17
Tab. 2. Parametry algoritmu MPEG	20
Tab. 3. Formáty obrazu pro video kodéry standardů H.261 a H.263.	24