

## Videokonference s vysokou kvalitou

Eva Hladká, Petr Holub, FI a ÚVT MU

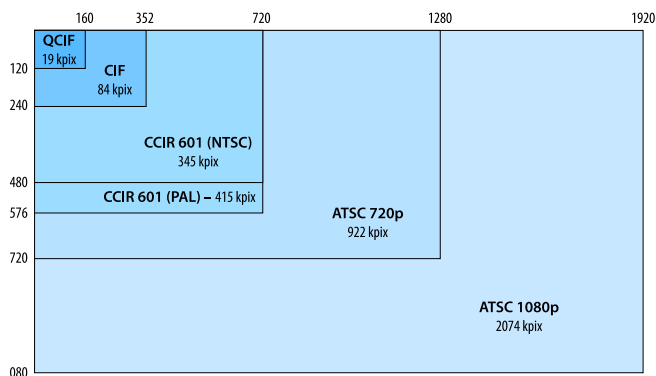
Videokonference se staly jednou z běžně využívaných technologií pro komunikaci jedinců i týmů. V řadě kanceláří se kromě telefonu a počítače nachází též videokonferenční zařízení nebo doplňky umožňující jako videokonferenční zařízení využívat počítač. Navzdory této skutečnosti, a nebo spíše právě kvůli ní, se i v oblasti videokonferencí bádá, vyvíjí a testuje. Zkoušejí se nové přístupy a vyšší kvalita obrazu a zvuku. Tento vývoj je na jedné straně urychlován rozvojem síťové infrastruktury a dostupností šířky použitelného pásma, na druhé straně jej vyžadují uživatelé, pro které je ergonomie a co největší přiblížení k realitě u videokonferencí velmi důležité. Přes veškerý pokrok je totiž videokonferenční komunikace pro zúčastněné namáhavější než běžná schůzka. Proto se v tomto příspěvku zaměříme na videokonference s vysokou kvalitou videa a zvuku. Je to vývojový trend, který se na MU za podpory VZ *Optická síť národního výzkumu a její nové aplikace* úspěšně rozvíjí.

### Video

Na tomto místě si čtenáři dovolueme připomenout několik základních parametrů, které ovlivňují výslednou kvalitu videokonference. Jedná se o kvalitu videa, šířku datového toku a zpoždění.

**Kvalita videa:** Použijeme-li pro videokonferenci zde již mnohokrát popsané Mbone Tools, potom malý obrázek v základním menu (QCIF) má  $180 \times 144$  bodů, zvětšený obrázek (CIF) má  $360 \times 288$  bodů. Pro srovnání - televizní přenos v normě PAL má  $720 \times 576$  bodů<sup>1</sup>. Mluvíme-li o videu s vysokým rozlišením, míníme tím v rámci tohoto článku některý z rodiny formátů dle standardu HDTV, kde snímek má rozlišení buď  $1280 \times 720$  nebo  $1920 \times 1080$  bodů. Pro větší názornost jsou jednotlivé formáty srovnány na obrázku 1.

<sup>1</sup>Technicky vzato je rozlišení  $720 \times 576$  platné pouze pro zařízení s obdélníkovými body jako jsou například televize. V případě zařízení se čtvercovými body, např. počítačové obrazovky, je rozlišení  $768 \times 576$  bodů.



Obrázek 1: Porovnání velikosti video obrazu pro různé formáty.

Celkovou kvalitu videa samozřejmě neovlivňuje pouze jeho velikost, ale i další faktory, jako barevná hloubka obrazu či použití ztrátové komprese. Zatímco běžné počítačové grafické karty pracují při přehrávání videa s 8 bity na barevný kanál, profesionální nasazení používá nejméně 10 bitů, což při třech barevných rovinách rozšiřuje barevnou škálu  $64 \times$ . Ztrátová komprese zejména při použití vysokého kompresního poměru často vede na různé artefakty v obraze: v lepším případě je to jen ztráta ostrosti detailů, v horším pak na posterizaci obrazu (rozpad obrazu do čtvercových bloků).

**Šířka datového toku:** V předchozím odstavci byla uvedena rozlišení základních videoformátů. Abychom měli představu o tom, kolik dat je třeba přenášet, chybí další parametr, a tím je počet obrázků za časovou jednotku, typicky sekundu - odtud anglická zkratka fps, neboli frames per second. Tento parametr ovlivňuje plynulost pohybu v přenášeném videu. U statické snímávané scény je možné počet snímků omezovat často až k 1 fps, pro plynulost pohybu u běžně komunikujících osob postačí 5-15 fps, ale pro skutečně plynulý pohyb dynamické scény je třeba přenášet 30 fps nebo i více. Výsledný datový tok pro HD formát v rozlišení  $1920 \times 1080$  dle HDTV normy se 60 prokládanými snímky za sekundu a 10 bity na barevný kanál je celkem  $1,5 \text{ Gb/s}^2$ . Na první pohled se může zdát, že pře-

<sup>2</sup>Pokud by laskavému čtenáři nevycházel výpočet potřebného datového toku z výše uvedených hodnot, pak je na správné stopě. V rámci výkladu jsme se dopustili dvou zjednodušení - skutečné HDTV rozlišení včetně tzv. ma-

nos takového množství dat je nereálný. Současné experimentálně-provozní sítě však mohou přenášet až desítky Gigabitů. Navíc, není-li striktně požadována maximální kvalita obrazu, minimální zpoždění a nezávislost jednotlivých snímků, lze videodata efektivně komprimovat. Z jednoho nekomprimovaného streamu lze získat například 25 Mb/s stream komprimovaný ve formátu HDV.

**Zpoždění:** Základní vlastností přirozené komunikace je zdánlivě okamžitá odezva. U videokonferencí je potřeba počítat se zpožděním (latencí), které vzniká zpracováním videosignálu u vysílající stanice, dobou potřebnou pro přenos dat sítě ke koncové stanici a zpracováním a zobrazením dat na koncové stanici. , viz obr. 2.

Počítáme-li s objemem dat 1,5 Gb/s, potom zpracování a přenos daného objemu bude vyžadovat netriviální kapacity a čas. Čím ovšem bude delší zpoždění, tím méně přirozená a tedy kvalitní bude komunikace. Lidské smysly (a ty jsou zde podstatné, protože koncovým uživatelem je člověk, nikoliv stroj) mají různou schopnost zaznamenat zpoždění. Při komunikaci je zapojen zrak i sluch a tím citlivějším a tedy směrodatným je sluch. Schopnost zvukové synchronizace se dá u člověka vytrénovat, ne nadarmo dosahují v tomto ohledu vynikajících časů například hudebníci hrající v komorních orchestrech, kde se přesnost synchronizace udává až kolem 5 ms. Běžně se doporučuje pro kvalitní komunikaci nepřekročit hranici 100 ms a proto je nutno minimalizovat časy nutné na zpracování obrazu na obou koncích. Za předpokladu, že používané sítě mají dostatečnou propustnost, je možné ušetřit čas odstraněním komprese. Zvláště pro přenosy na vzdálenosti v řádu deseti a více tisíc kilome-

---

zacích rádků je  $2200 \times 1125$ . Dále se pro snížení datového toku využívá menší citlivosti lidského oka na barvy než na jas, a proto je barevný prostor YCrCb vzorkován 4:2:2, což sníží datový tok na  $2/3$ . Správný výpočet je tedy

$$2200 \times 1125 \times 30 \times 30 \times 2/3 = 1,485 \text{ Gb/s.}$$

Tento formát se také označuje jako HD-SDI specifikovaný v normě SMPTE 292M [1].

trů<sup>3</sup>, kde i rychlost přenosu dat není zanedbatelnou položkou v celkovém zpoždění.

## Zvuk

Zatím jsme explicitně nezmínili zvuk, i když jeho kvalita je pro úspěšnost a ergonomii videokonference podstatná. U videokonferencí s vysokou kvalitou obrazu je rozhodující právě rozlišení videa, kvalita zvuku však nesmí komunikaci negativně ovlivnit.

K HDTV obrazu může být připojeno prakticky libovolné audio, takže lze využít například nejvyšší kvality dle standardu High-Definition Multimedia Interface (HDMI) [2], tj. nízkolatenční nekomprimovaný zvuk s 24 bitovým kvantováním, vzorkováním 192 kHz a s 8 (7.1) kanály, což dává 36,8 Mb/s. I v této značně maximalistické kvalitě však představuje zvuk pouze malý zlomek přenášených dat v porovnání s nekomprimovaným HDTV videem.

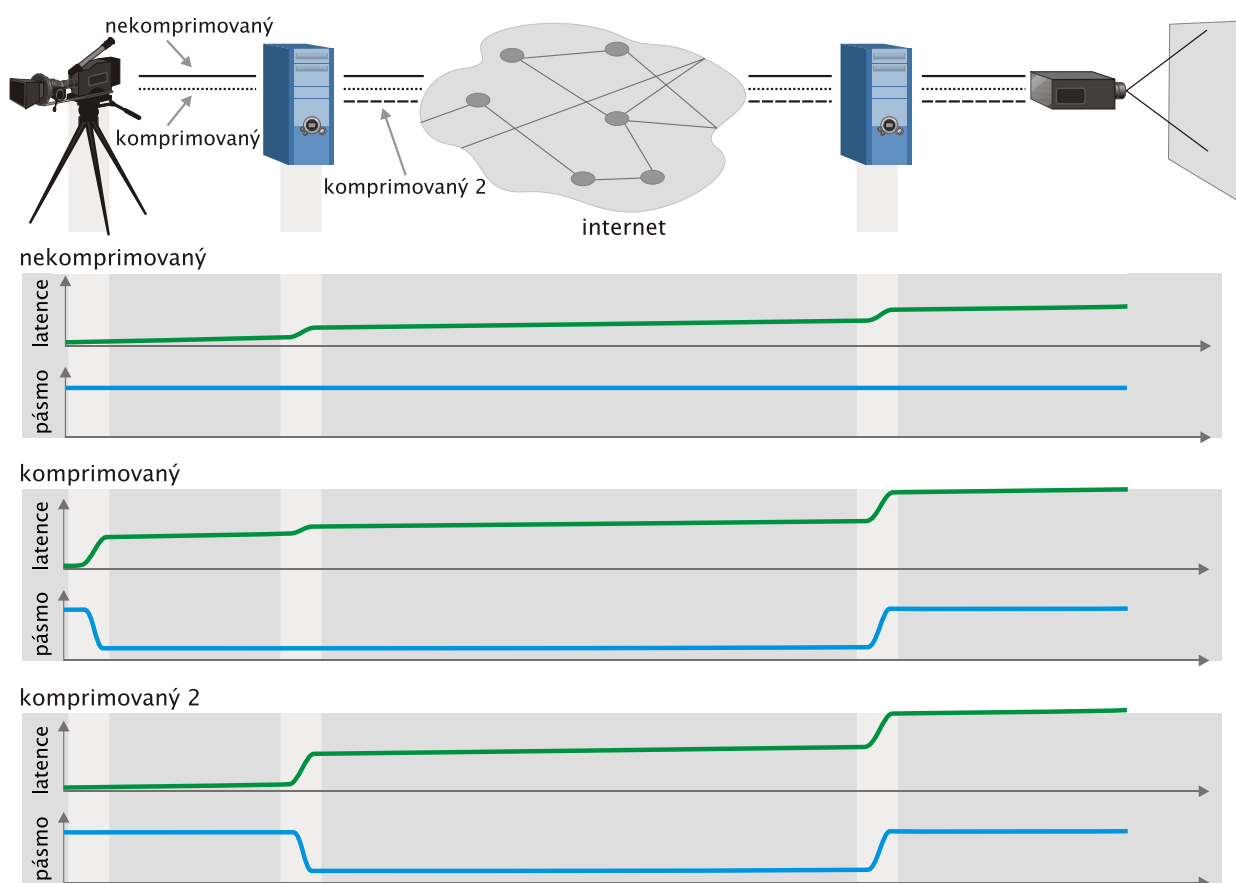
Vezmeme-li v potaz skutečnost, že lidský sluch je na zpoždění mnohem vnímavější než zrak, nabízí se na první pohled lákavá možnost posílat zvuk nekomprimovaný s co nejnižší latencí, zatímco obraz by se extenzivně komprimoval pro zmenšení nároků na přenosovou kapacitu. Zde však narazíme na další záludnost lidského vnímání: podobně jako je lidský sluch citlivý na celkové zpoždění, je zrak citlivý na synchronizaci mezi zvukem a obrazem - v angličtině se běžně užívá termínu "lip synchronization", tedy synchronizace na rty. Hranice citlivosti na desynchronizaci obrazu a zvuku se pohybuje opět přibližně kolem 100 ms a nekomprimované video se z tohoto pohledu opět ukazuje jako vhodný formát.

## Komunikační schéma

Zatím jsme nezmínili mezi kolika účastníky může HD videokonference probíhat, tedy jaké komunikační schéma lze použít. Základním omezujícím faktem je, že koncové místo vysílá

---

<sup>3</sup>Při rychlosti světla 300 000 km/s urazí světlo vzdálenost 10 000 km za 33 ms. Je však nutno vzít v potaz, že světlo se v optickém kabelu nešíří stejně rychle jako ve vakuu, ale 1,5-násobně pomaleji, a tudíž stejnou trasu urazí za zhruba 50 ms.



Obrázek 2: Nárůst latence u videopřenosu. Vodorovná osa na obrázku odpovídá poloze videa v řetězci zpracování, jak je uveden v horní části obrázku. Rozdíl mezi komprimovaným videem a komprimovaným videem 2 spočívá v tom, že v prvním případě probíhá komprese přímo v kaměře, kdežto ve druhém případě jde video z kamery nekomprimované a ke kompresi dochází teprve v počítači.

1,5 Gb/s a nejméně stejný objem dat přijímá. Maximální počet účastníků tedy omezuje kapacita koncové linky. Dalším problémem je replikace. Pokud bude počet účastníků větší než dva, je třeba vysílaná data replikovat tak, aby je dostala všechna přijímající místa. Řešení známá z videokonferencí v běžné kvalitě zde selhávají. Skupinová komunikace v podobě multicastu není adaptována na potřebné objemy dat zejména v prostředí heterogenních sítí stejně tak jako MCU jednotky známé z H.323 videokonferencí.

Problém komunikace mezi více účastníky lze řešit za použití speciálně upravených vysoce výkonných reflektorů multimediálních datových toků [3] nebo přímé multiplikace signálu na optické vrstvě pomocí splitterů. Výhoda optických

splitterů je v tom, že nepřidávají zpoždění a jsou takřka nezávislé na přenosové rychlosti (a dokonce mohou být širokospektrální a dělit více vlnových délek najednou), problémem je však jejich malá flexibilita a zatím pouze experimentální dostupnost.

### Závěrem

Zejména výše uvedené problémy s vícebodovou distribucí zatím využití rozsáhlejších videokonferencí na bázi nekomprimovaného videa značně omezují. Prvními vlaštovkami v tomto směru mohou být demonstrace na workshopu iGrid 2005 popsané v samostatném článku v tomto čísle Zpravodaje, kde dva týmy nezávisle na sobě tyto technologie demonstrovaly.

HD videokonference a jejich další zlepšení posunují oblast prostředí pro virtuální spolupráci ke stále reálnějšímu a přirozenějšímu vjemu účastníků. V brzké době umožní přenosy z míst, kde je kvalita detailu obrazu velmi důležitá, např. z operačních sálů do poslucháren. Pro nás představují mnoho výzev a ukazují mnoho problémů, které je třeba řešit. Na rutinní využití těchto technologií si ještě budeme muset nějaký čas počkat, ale do budoucna je třeba s nimi vážně počítat.

## Literatura

- [1] Society of Motion Picture and Television Engineers. *Bit-Serial Digital Interface for High-Definition Television Systems*. SMPTE 292M-1998.
- [2] *High-Definition Multimedia Interface (HDMI)*. <http://www.hdmi.org/>
- [3] E. Hladká, P. Holub. „Zrcadla v počítačové síti.“ *Zpravodaj ÚVT MU*. ISSN 1212-0901, 2002, roč. 12, č. 5, s. 7-10. □