

Využití holofonního prostorového zvuku pro účel rozšířené reality

1 – Projektové zhodnocení vývoje nové aplikace – čas, náklady atd.

Vývoj aplikace nelze hodnotit dokonale, protože je stále ve fázi prototypu. Projekt byl navržen tak, aby nevyžadoval nákup specializovaného HW (jen mobil a sluchátka) a k jeho realizaci nebyl použit placený software. Můžeme tedy mluvit jen o nákladech na práci.

Asi dva měsíce jsem získával informace a hledal vhodný toolchain pro zpracování zvukových vzorků. Další dva měsíce jsem programoval aplikaci samotnou. Největší překážkou byl fakt, že jsem nikdy předtím aplikaci pro Android nedělal a tak jsem musel vyzkoušet několik frameworků a vývojových prostředí. Po seznámení s SDK už bylo vytvoření prototypu otázkou několika málo dní. Vývojář s předchozími zkušenostmi by jistě postupoval výrazně rychleji. Také je nutné zdůraznit, že jsem prototyp vyvíjel paralelně se školou a zaměstnáním. Odhaduji tedy, že šlo zhruba o desítky hodin čistého času.

2 – Jaké může být zpoždění zobrazované reality u pohybujícího se objektu (mili, mikro sekundy) k vykreslení zobrazovací matice v návaznosti na rychlost tohoto objektu?

Vycházím z předpokladu, že lidské oko vnímá maximálně zhruba 30 snímků za vteřinu. Tomu přizpůsobím i kameru a zobrazovací matici. Pokud předpokládám, že kamera i zobrazovací matice snímají synchronizovaně, bude možné vypočítaný obraz AR zobrazit nejdříve za asi 33 milisekund. To je doba, kterou mohu využít k vypočítání a vykreslení obrazu AR bez výrazného pocitu zpoždění.

Nabízí se ale i varianta, kdy máme kameru s velmi citlivým a tedy rychlým snímačem, což mi umožní získat snímek o něco později a tedy aktuálnější vzhledem k následujícímu snímku zobrazenému na monitoru. Pokud pak odpovídající měrou zkrátím čas na zpracování AR, budu mít nepatrně aktuálnější obraz.

Řádově lze tedy odpovědět, že zpoždění zobrazované reality by nemělo být delší než 33 ms, ale může mít smysl jeho snižování až zhruba ke 3 ms.

3 – Jaká je úroveň je zkreslení při „downsamplu“ na 22.05 kHz – viz. kvantizační zkreslení v PCM.

Ze Shannon-Nyquistova vzorkovacího teorému vyplývá, že při samplovací frekvenci 22 kHz budeme schopni zaznamenat frekvence do 11 kHz. Pokud budeme downsamplovat existující záznam z 44.1 kHz na 22.05 kHz, tedy přesně na poloviční samplovací frekvenci, nehrozí ztráta informace při aliasingu.

Lokalizace zdroje zvuku lidským sluchem pracuje převážně s frekvencemi do 5 kHz, výjimečně do 10 kHz (elevace). Samplovací frekvence 22 kHz by měla tedy být pro potřeby mé práce dostatečná. Přesto jsem použil vyšší (lepší) samplovací frekvenci 44.1 kHz, protože je dnes poměrně rozšířená (audio CD,...) a protože HRIR data jsem získal právě v tomto vzorkování.

