

Univerzita Komenského
Fakulta matematiky, fyziky a informatiky
Katedra informatiky

Radoslav Buranský

Návrh, implementácia a ohodnotenie
zobrazenia informácií do zvukového
priestoru
(diplomová práca)

Vedúci diplomovej práce: doc. RNDr. Roman Ďurikovič, PhD.
Odbor: Programové a počítačové systémy

Bratislava, Máj, 2007

Čestne vyhlasujem, že som túto diplomovú prácu
vypracoval samostatne s použitím citovaných zdrojov.

.....
Radoslav Buranský

1 Abstrakt

V našej práci sme sa snažili navrhnúť zobrazenie informácií obsiahnutých v HTML dokumentoch do trojrozmerného zvukového priestoru. Sústredili sme sa na praktické použitie bez nutnosti použitia špeciálnych výstupných zariadení. Vychádzali sme z najnovších poznatkov o vnímaní priestorového zvuku. Pre dosiahnutie čo najkvalitnejšieho riešenia sme nápady konzultovali s nevidiacimi. Výsledný návrh sme implementovali, a ten následne testovaním ohodnotili. Z hodnotenia vyplynula použiteľnosť navrhnutého riešenia, a aj jeho výhody oproti bežne používaným zvukovým prehliadačom.

2 Obsah

1	Abstrakt	4
2	Obsah	5
3	Úvod	7
4	Cieľ a motivácia	8
5	Použité metódy	9
5.1	<i>VIPER</i>	9
5.2	<i>Vstupné dáta</i>	9
5.3	<i>Rečový syntežátor</i>	10
5.4	<i>Zvuková banka</i>	10
5.5	<i>Priestorový zvuk</i>	11
5.5.1	HRTF	11
5.5.2	Odrazy zvuku	13
5.6	<i>Súčasné vnímanie viacerých zvukov</i>	13
5.7	<i>Navigácia</i>	14
5.7.1	DAHNI.....	15
5.7.2	Dynamic Soundscape	16
5.7.3	A 3D Audio Only Interactive Web Browser	17
5.7.4	Jaws.....	22
6	Návrh zobrazenia	24
7	Implementácia	34
7.1	<i>Festival</i>	34
7.2	<i>OpenAL</i>	37
7.3	<i>ALUT</i>	38

7.4	<i>Microsoft Internet Explorer</i>	38
7.5	<i>B4BLib</i>	39
7.6	<i>B4BReader1Lib</i>	40
7.7	<i>B4B</i>	41
7.8	<i>Architektúra</i>	42
8	Testovanie	43
8.1	<i>Úloha č. 1</i>	44
8.2	<i>Úloha č. 2</i>	46
8.3	<i>Úloha č. 3</i>	49
8.4	<i>Výsledky</i>	50
9	Prílohy	51
9.1	<i>Návod na použitie</i>	51
9.1.1	<i>Inštalácia</i>	51
9.1.2	<i>Užívateľské rozhranie</i>	51
10	Referencie	54

3 Úvod

Schopnosť pracovať s počítačom sa považuje za samozrejmosť. Ľudia so zrakovým hendikepom však narážajú na bariéru, ktorá im v práci bráni, aj keby nemusela. Mnohé z programov sa totižto dajú obsluhovať aj bez zrakového vnemu. Stačí mať pomocný program, ktorý obsah obrazovky prenáša do podoby vnímateľnej pre nevidiacich, a tou je zvuk. Obmedzenia pri práci s počítačom nemusia mať len ľudia s poškodeným zrakom. Môže ísť napríklad o vodičov motorových vozidiel, ktorí potrebujú popri jazde nájsť najbližší hotel. Inou možnosťou aplikácie je použitie v mobilných zariadeniach, kde si užívateľ môže popri inej činnosti prezerat' internetové stránky.

Riešení existuje niekoľko, ale aj tie najkvalitnejšie, z bežne dostupných, využívajú len monofónny zvuk. Využitím priestorového zvuku sa dá dosiahnuť oveľa bohatšie podanie informácie. V minulosti bolo takéto riešenie finančne náročné, avšak dnes si môže viacnásobnú zvukovú zostavu zaobstarat' takmer každý. Navyše priestorový vnem je možné získať aj použitím bežných slúchadiel.

4 Cieľ a motivácia

Cieľom našej práce je navrhnuť, implementovať a ohodnotiť zobrazenie obsahu HTML dokumentov do zvukového priestoru. Výsledok bude použitý v európskom projekte VIPER (angl. Visual Attention Processes for Electronic Content Restructuring), ktorého cieľom bude poskytnúť ľuďom s obmedzeným zrakovým vnemom možnosť prezerania HTML stránok. Dôraz je kladený na praktické použitie pre širokú verejnosť. Program preto nesmie mať vysoké výpočtové ani technické požiadavky. Z tohto dôvodu bude jeho súčasťou implementácia metód, ktoré umožňujú priestorový vnem použitím slúchadiel. Výsledky budú konzultované s nevidiacimi, aby sme dosiahli čo najkvalitnejšie riešenie.

5 Použité metódy

5.1 VIPER

Viper je rozsiahly projekt, ktorý rieši prezeranie HTML stránok zo všetkých pohľadov. Spolupracuje na ňom šesť tímov zriadených na európskych univerzitách. Ich výskum sa sústreďuje na niekoľko oblastí:

- Používanie internetového prehliadača vo všeobecnosti.
- Zobrazenie HTML stránok na viacero výstupných zariadení ako sú zvukový výstup, hmatový displej a haptické zariadenie. Haptický ovládač je vstupno-výstupné zariadenie, ktorým sa dá po uchopení rukou pohybovať vo všetkých troch smeroch. Informácie o pohybe sa prenášajú do počítača, ktorý ich spracuje a ako reakciu môže vyslať príkaz ovládaču, aby spätne pôsobil silou na ruku používateľa. Vďaka tejto funkcii je možné približne zistiť obrysy nejakého trojrozmerného predmetu.
- Anotácia dokumentov. Cieľom je určiť, ktoré prvky sú významovo a vizuálne dôležité.
- Testovanie.

Časový plán počíta s troma rokmi práce. V čase implementácie našej práce ešte neboli dokončené potrebné moduly, a teda ich nebolo možné v našej práci využiť. Avšak nás to nijak významne neobmedzilo.

5.2 Vstupné dáta

Ako už bolo spomenuté, náš program bude súčasťou projektu VIPER, a teda vstupné dáta budú získané od niektorého z jeho modulov. Keďže VIPER je zatiaľ v štádiu návrhu, bližšie informácie o formáte vstupných dát zatiaľ nie sú k dispozícii. S určitosťou môžeme povedať, že budú vychádzať z dátovej štruktúry HTML. Dočasný dátový model bude vychádzať z týchto poznatkov

a bude navrhnutý dostatočne flexibilne, aby sa dal bez ťažkostí transformovať na budúci skutočný dátový model projektu VIPER, ktorý naše dáta doplní o informáciu o vizuálnej a významovej atraktivnosti. Táto vlastnosť je jedna z kľúčových v projekte VIPER. V implementácii používame ako modul Microsoft Internet Explorer, ktorý zabezpečuje syntaktickú a sémantickú analýzu dokumentu.

5.3 Rečový syntezátor

Automatizovaný prevod textu na zvuk sa bude vykonávať pomocou rečového syntezátora. Ide o matematický model hlasiviek, pričom neobsahuje žiadne vopred pripravené zvukové záznamy. Samotný rečový syntezátor je vytvorený všeobecne, aby vedel simulovať ľubovoľný jazyk a farbu hlasu. Pre jednotlivé jazyky sú vytvorené lexikóny, ktoré obsahujú intonačné pravidlá špecifické pre daný jazyk. Použijeme voľne dostupný syntezátor Festival a spolu s ním niekoľko anglických lexikónov. Lexikón sú dáta popisujúce vlastnosti hlasu ako je výška, farba, intonácia a podobne. Zvyčajne je pre každý syntezovaný hlas práve jeden lexikón.

5.4 Zvuková banka

Popri reči budeme využívať rôzne zvuky neobsahujúce reč (angl. earcons). Takéto zvuky majú oproti reči mnoho výhod. Informujú o udalosti, ktorá práve nastala, pričom nerušia pozornosť od práve čítaného textu. Sú krátke a tým urýchľujú prácu. Nerečový zvuk môže byť abstraktný zvuk, tón hudobného nástroja, alebo zvuk s jasným významom (zvuk písacieho stroja, krčenie papiera, a pod.). Umiestnenie nerečového zvuku v priestore môže podať doplňujúcu informáciu (napríklad pozíciu v čítanom texte). Navyše sa môžu dynamicky meniť niektoré jeho vlastnosti ako napríklad intenzita alebo výška tónu.

5.5 Priestorový zvuk

Umiestnenie monofónneho zvukového signálu do priestoru je dôležitá vlastnosť nášho projektu. Dosiahne sa tým obohatenie zvukovej informácie. Pod umiestnením zvuku do priestoru rozumieme simuláciu javov prítomných v normálnych akustických podmienkach. Medzi ne patrí pokles hlasitosti s rastúcou vzdialenosťou od zdroja zvuku a odrazy od stien a prekážok. Pri použití viackanálového výstupného zvukového zariadenia sú tieto dve zákonitosti postačujúce pre vierohodný priestorový vnem. Kvalita závisí už len od množstva a rozmiestnenia reproduktorov okolo poslucháča.

Náš program má za cieľ dosiahnuť uspokojivý priestorový zvuk aj pomocou stereofónnych slúchadiel. V tomto prípade je nutné modelovať akustické vlastnosti hlavy. Pod akustickými vlastnosťami hlavy rozumieme:

- Interaural Intensity Difference (ITD) – zvuk vychádzajúci zprava od poslucháča bude znieť hlasnejšie v pravom uchu ako v ľavom.
- Interaural Time Difference (ITD) – zvuk vychádzajúci zprava dorazí do pravého ucha skôr ako do ľavého.
- Frekvenčná charakteristika – hlava a vonkajšie uši pôsobia ako zvukové filtre.

Dve najdôležitejšie vlastnosti sú rozdiely v čase pre nízke frekvencie a rozdiely v intenzite pre vysoké frekvencie. Pri viackanálovom zariadení to nutné nebolo, keďže zvuk vychádzajúci z príslušného reproduktora bol ovplyvnený hlavou poslucháča skôr ako bol uchom vnímaný.

5.5.1 HRTF

Funkcia popisujúca akustické vlastnosti hlavy, pre zvuk prichádzajúci z daného smeru, sa nazýva HRTF (angl. Head Related Transfer Function).

Smer zvuku sa udáva dvojicou uhlov. Prvý z nich sa označuje azimut (horizontálna rovina) a druhý ako výška (vertikálna rovina).

Výstupom HRTF pre daný monofónny zvuk, azimut a výšku je stereofónny zvukový signál. HRTF sa získava meraním v laboratórnych podmienkach pre konkrétneho jedinca. Meranie sa vykonáva najmä pre získanie frekvenčnej charakteristiky. Tá má kľúčovú úlohu pri lokalizácii zvukov. Každá osoba má túto funkciu inú, preto HRTF delíme na individualizované a neindividualizované. Neindividualizovaná HRTF je tá, ktorá bola získaná pre iného jedinca než je poslucháč.

Väčšina poslucháčov je schopná získať základné informácie o smere pre zdroj zvuku aj pomocou neindividualizovanej HRTF. Najľahšie sa určuje azimut. Naopak pri určovaní výšky zvuku sú časté problémy.

Zaujímavými typmi nejasností sú predozadné (angl. front-back) a horno-dolné (angl. top-down). V prvom prípade má poslucháč pocit, že zvuk, ktorý v skutočnosti ide spredu, má zdroj vzadu (alebo naopak). Podobne je to aj s pomýlením hornej a dolnej hemisféry. Predpokladá sa, že medzi príčinami tohto paradoxu je vžitý predpoklad človeka, že to, čo nevidí, musí byť za ním. Riešenie pre tento problém zatiaľ nie je, ukazuje sa však, že vhodné pridanie ozvien môže pravdepodobnosť nejasností znížiť.

Ďalším problémom je, že potrebujeme mať HRTF pre ľubovoľný smer, ibaže experimentálnym meraním sa získa len pre konečný počet smerov. Vychádzajúc zo zistení viacerých autorov je postačujúce použiť lineárnu interpoláciu.

Samotné HRTF funkcie sú dajú voľne získať od americkej univerzity MIT. Obsahujú dáta pre dva modely vonkajšieho ucha a aj frekvenčné charakteristiky štyroch populárnych slúchadiel, pre čo najvernejšiu simuláciu. Merania boli vykonané pre 710 rôznych smerov.

Priama implementácia HRTF v našom projekte nie je nutná, keďže v súčasnosti už aj cenovo dostupné výstupné zvukové zariadenia ju obsahujú.

5.5.2 Odrazy zvuku

V bežných akustických podmienkach každý zvuk, ktorý počujeme, má ozvenu. Jej intenzita, dĺžka, a iné vlastnosti nám dávajú informáciu o vzdialenosti od zdroja zvuku a vďaka nej vnímame veľkosť priestoru. Výsledky výskumov ukazujú, že ozvena znižuje presnosť určovania smeru zvuku. Na druhej strane pomáha pri určovaní vzdialenosti, a zvuky znejú príjemnejšie a prirodzenejšie. Zaujímavou vlastnosťou je schopnosť človeka prispôbovať sa danej miestnosti. Čím dlhšie sa v nej nachádza, tým presnejšie vie určiť smer zvuku. Dôležitým faktorom pre zváženie použitia ozvien je aj výpočtová náročnosť. Kvalitná simulácia v reálnom čase nie je na dnešných bežných počítačoch možná.

5.6 Súčasné vnímanie viacerých zvukov

Použitie prekrývajúcich sa zvukov môže podstatne urýchliť získavanie informácií. Treba však brať na vedomie, že človek nie je schopný plnohodnotne vnímať príliš veľa súčasne znejúcich zvukov. Známym javom je “cocktail party effect”. Je to schopnosť človeka sústrediť sa spomedzi viacero zvukov na jeden z nich. Jednotlivé zvuky sa však musia dostatočne líšiť v parametroch ako sú smer, farba, intenzita. Dôležité vlastnosti sú aj spojitosť zvuku a jeho vzťah s vizuálnym objektom. Je ale dokázané, že vnímanie ostatných zvukov, na ktoré poslucháč práve nie je sústredený, je veľmi malá.

Autori výskumu Brungart a Simpson sa snažili nájsť optimálne rozmiestnenie siedmich rečníkov vo zvukovom priestore tak, aby poslucháč vedel čo najľahšie určiť, ktorý z nich povedal vopred zadané kľúčové slovo. Poslucháč sa potom na tohto rečníka sústredil. Z výsledkov síce vyplýva, že je možné vnímať paralelné zvuky, ale cieľom bolo len zachytiť vopred známe slovné spojenie. Pre našu prácu tento prístup využiteľný nie je, keďže náš

užívateľ nevie vopred obsah prezeraného dokumentu a jeho záujmom nie je zistiť, či sa jedno slovo v dokumente nachádza, ale zschytiť čo najviac z nich.

V inej práci sa autori na problém súčasne znejúcich zvukov pozreli inak. Myšlienka simulácie budovy s miestnosťami, v ktorých prebiehajú rozhovory, je zaujímavým využitím spomínaného cocktail party efektu. V každej miestnosti sa vedie diskusia na nejakú tému. Poslucháč kráča po chodbe a počuje viacero rozhovorov súčasne. Vojde iba do tej miestnosti, z ktorej počuje rozhovor o tom, čo ho zaujíma. Je komplikované algoritmicky simulovať diskusiu o obsahu prezeraného dokumentu, a preto ani tento prístup nie je pre nás zaujímavý.

Schopnosť sústrediť sa na určitý zvuk spomedzi iných, sa dá využiť pri rýchlom hľadaní informácie v texte. Celý text sa rozdelí na niekoľko častí, pričom každú časť číta iný rečník. Túto myšlienku použili autori projektu *Dynamic Soundscape*, ktorý bude bližšie popísaný v jednej z nasledujúcich kapitol. Rečníci sú umiestnení v rovine okolo hlavy. Poslucháč, na to aby našiel hľadanú informáciu, nemusí robiť nič, len postupne sa sústrediť na každého rečníka.

5.7 Navigácia

Pod navigáciou rozumieme metódy prezerania vstupných dát vo zvukovom priestore. Navigácia je úzko spätá s ovládacími zariadeniami, ktoré užívateľ používa pri práci s programom. VIPER bude využívať najmä klávesnicu a haptické výstupné zariadenie Phantom. V našej práci budeme využívať iba klávesnicu, keďže Phantom je drahé zariadenie, a našim cieľom je vytvoriť riešenie pre čo najviac užívateľov.

Súčasťou navigácie je aj voľba poradia dátových elementov, v ktorom budú programom zobrazené. Je vhodné aby bol k dôležitým informáciám rýchly prístup. Takisto záleží na rozmiestnení v priestore, výbere vhodných earconov pre rôzne udalosti a podobne. V nasledujúcich kapitolách zhrnieme

základné myšlienky viacerých spôsobov navigácie, z ktorých budeme pri našom návrhu čerpať.

5.7.1 DAHNI

“Demonstrator of the ACCESS Hypermedia Non-visual Interface” (DAHNI) je návrh rozhrania programu na prezeranie hypermediálneho obsahu pre nevidiacich. Využíva syntezovanú reč, digitálne nahrávky niektorých slov, nerečové zvuky, Brailleovu obrazovku pre výstup znakov, klávesnicu, pákový ovládač a veľkú dotykovú obrazovku.

Vstupné informácie sú rozdelené do jednotiek. Pod jednotkou sa rozumie celý uzol, odstavec, veta, slovo alebo písmeno. Umožní sa tým práca s údajmi na rôznej úrovni detailu. Pre získanie prehľadu o obsahu, DAHNI poskytuje rýchlu navigáciu po všetkých odkazoch, taktiež informuje o názve a nadpisoch pre daný uzol. Pre zrozumiteľnejšiu prácu s program používa dva typy hlasov. Jeden, syntetizovaný, číta obsah údajov, a druhý, vopred nahraný, číta názvy príkazov programu. Rozlíši sa tým obsah od užívateľského rozhrania. Príkazy sú usporiadané do tvaru písmena H. Užívateľ si pomocou niektorého zo vstupných zariadení volí aktívny príkaz, ktorý chce vykonať. Názov príkazu zaznie zo zvukového zariadenia. Medzi príkazmi sú napríklad: predchádzajúci, nasledujúci, stop, opakuj, najdi, nasleduj odkaz a podobne.

Ako autori poznamenávajú, kľúčová je práca s odkazmi. Dva najväčšie rozdiely medzi vizuálnym a nevizuálnym rozhraním sú v spôsobe, akým je odkaz prezentovaný, a akým sa odkaz nasleduje. Pre prezentáciu navrhujú použitie iného hlasu ako na čítanie ostatného obsahu. Inou možnosťou je nerečový zvuk, ktorý zaznie pred prečítaním odkazu. Problém s nasledovaním odkazu riešia tak, že odkaz, ktorý bol prečítaný ako posledný, sa označí ako aktívny. Aktívnym zostáva až do okamihu kým sa neprečíta iný odkaz. Ak sa užívateľ rozhodne odkaz nasledovať, zvolí si príkaz „nasleduj odkaz“.

Prieskum ukázal, že nie každému vyhovoval zvuk zvoleného earconu, a preto by program mal poskytovať možnosť výberu z viacerých možností. Podobne to je aj so spôsobom, akým sa earcon prezentuje. Najobľúbenejším je keď zvuk znie počas čítania odkazu.

Ďalším zistením bolo, že užívatelia nepovažujú za veľmi dôležitú informáciu o type odkazu, a teda či odkaz smeruje na text, obrázok, alebo iný typ dát. Postačovalo, keď sa to dozvedeli až po jeho nasledovaní.

Veľmi pozitívne bolo hodnotené použitie vopred nahraného hlasu pre čítanie príkazov. Rýchlosť čítania textu syntetizátorom 140 slov za minútu bola uspokojivá pre väčšinu užívateľov.

5.7.2 Dynamic Soundscape

Problém hľadania informácie vo zvukovej nahrávke sa snaží tento program riešiť využitím priestorového zvuku, schopnosti človeka sústrediť sa na jeden z viacerých súčasne znejúcich zvukov a vlastnosti ľudskej pamäte, ktorá si vie k informácii pamätať aj pozíciu v priestore. Použité metódy sú zaujímavé aj pre náš program.

Nahrávka je rozdelená na niekoľko častí a tie sú rovnomerne rozmiestnené v rovine okolo hlavy poslucháča. Jednotlivé časti sa dajú aktivovať a tým sa spustí ich prehrávanie. Súčasne môžu byť aktívne štyri časti. Tu sa využíva schopnosť sústrediť sa na jeden zvuk z viacerých pokiaľ sú jednotlivé zvuky dostatočne odlišené. V tomto prípade je rôzna ich pozícia v priestore.

Podstatné je, že ak užívateľ informáciu nájde, jeho pamäť si pozíciu v priestore podvedome zaznamená. V budúcnosti keď bude tú istú informáciu potrebovať, spomenie si kde v priestore sa nachádzala a aktivuje príslušnú časť nahrávky.

Autori spomínajú niekoľko problémov, ktoré sa ukázali počas používania programu: zabudnutie pozície kde sa informácia nachádza,

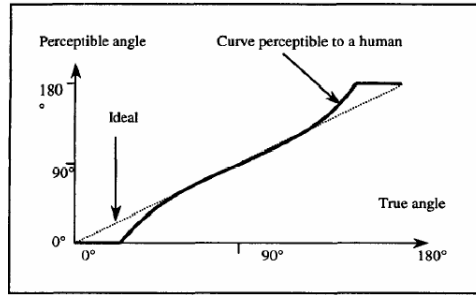
počúvanie viacerých zvukov súčasne, chyby pri určovaní polohy zvuku v priestore, rozlíšenie priestorovej pamäte a aktivovanie tej správnej časti nahrávky. Ako riešenie niektorých komplikácií navrhujú rozdelenie nahrávky na najviac 12 častí. Zároveň uvádzajú, že aktivovaná časť nahrávky by sa počas prehrávania nemala pohybovať. Zvýši sa tým presnosť zapamätania si správnej pozície v priestore.

Pozorovaním zistili, že rozmiestnenie zvukov v priestore pomáha prenosu informácii do dlhodobej pamäte. Je to vďaka vytvoreniu organizovanej priestorovej štruktúry z lineárneho zvukového záznamu.

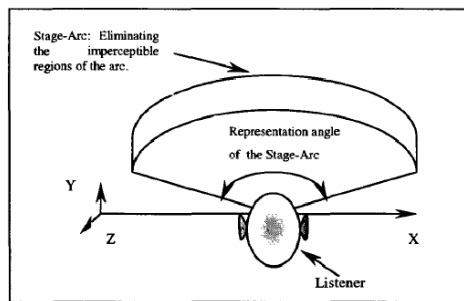
5.7.3 A 3D Audio Only Interactive Web Browser

Program, ktorý ako výstup využíval aj priestorový zvuk, bol navrhnutý pre spoločnosť Siemens. Spôsobmi navigácie nás pri návrhu inšpiroval v najväčšej miere. Vychádza z návrhov, ktoré sme spomínali v predchádzajúcich kapitolách. Je zameraný na praktické použitie a rieši mnoho základných problémov. Na simuláciu zvukového priestoru používa akustický model hlavy.

Autori sa rozhodli rozmiestňovať zvuky v horizontálnej rovine, keďže pri použití spomínaného modelu hlavy mali užívatelia problémy s určovaním výšky zvuku. Aby zabránili predozadným nejasnostiam, využívajú len prednú polrovinu. Zvuky sú umiestňované na okraj kruhového výseku so stredom v hlave poslucháča. Kvôli vlastnosti ľudského ucha (Obrázok 1), ktoré v okrajových bodoch nevníma azimut zvuku presne, netvorí kruhový výsek polkruh, ale je o čosi menší (Obrázok 2).

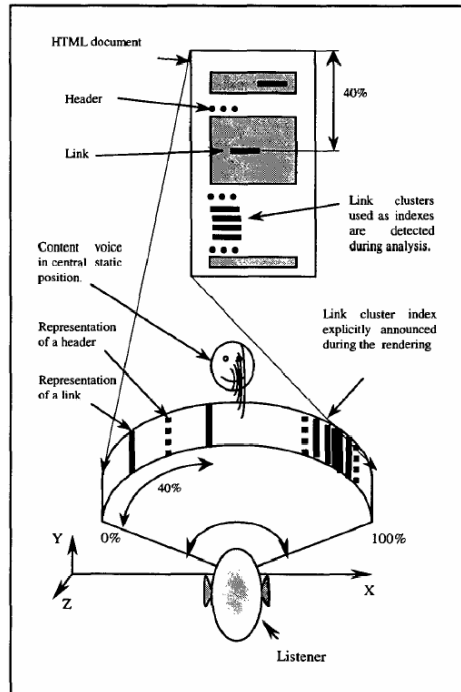


Obrázok 1: Krivka vnímania azimutu zvuku ľudským uchom.



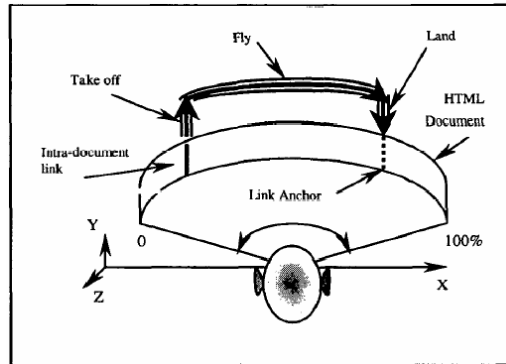
Obrázok 2: Kruhový výsek, na ktorý sú zvuky rozmiestňované.

Program používa tri rôzne syntetizované mužské hlasy. Jeden je určený na čítanie nadpisov, druhý číta obsah textov a tretí odkazy. Všetky tri sú umiestnené na kruhovom výseku priamo pred poslucháčom a svoju polohu nemenia. Štvrtý ženský hlas v pravidelných intervaloch oznamuje pozíciu v dokumente tým, že sa mení jeho poloha na výseku. Keď sa v čítanom texte nachádza odkaz, zaznie nerečový zvuk, ktorý odkaz reprezentuje (Obrázok 3).



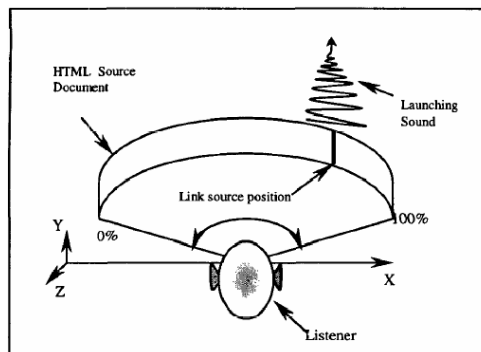
Obrázok 3: Rozmiestnenie zvukov.

Ako sme spomínali už skôr, základným problémom je prezentácia a nasledovanie odkazov. Odkaz môže byť vnútorný, odkazujúci na inú časť toho istého dokumentu, v ktorom sa nachádza, alebo vonkajší, ukazujúci na iný dokument. Autori článku uvádzajú, že pokiaľ sa tieto dva typy nerozlišia, nevidiaci užívatelia môžu ľahko stratiť orientáciu. Preto autori prehliadača navrhli nasledovné riešenie. Vnútorný odkaz bude používať tri typy zvukov: štartovací, letiaci a pristávací. Keď užívateľ aktivuje vnútorný odkaz, zaznie štartovací zvuk na príslušnej pozícii. Nasleduje letiaci zvuk, ktorý sa pohybuje smerom k miestu, kam odkaz smeroval. Nakoniec zaznie zvuk pristávací (Obrázok 4).

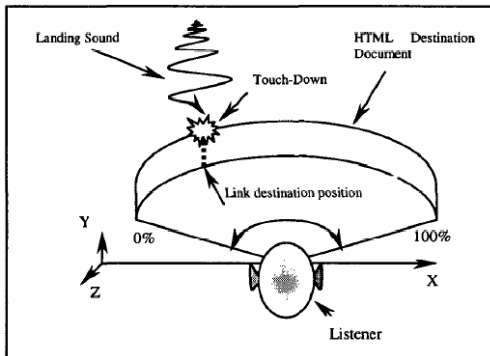


Obrázok 4: Vnútorný odkaz.

Pri vonkajšom odkaze sú zvuky dva: štartovací a pristávací. Sú odlišné od zvukov použitých pri vnútornom odkaze. Štartovací zvuk sa postupne vzdáľuje dávajúc najavo, že sa opúšťa práve prezeraný dokument (Obrázok 5). Pristávací zvuk používa opačný efekt približovania sa (Obrázok 6).

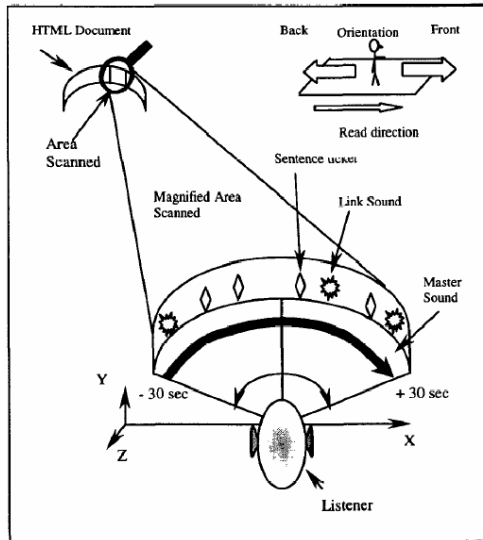


Obrázok 5: Štartovací zvuk pre vonkajší odkaz.



Obrázok 6: Pristávací zvuk pre vonkajší odkaz.

Pri vizuálnom prezeraní hypermediálneho dokumentu najprv získame prehľad o jeho obsahu rýchlym prezretím. Poskytnúť túto možnosť aj nevidiacim sa rozhodli autori nasledovným spôsobom (Obrázok 7). V časti dokumentu pred a za aktuálnou pozíciou sa nájdu vety a odkazy. Dĺžka spracovanej časti je udaná v sekundách, a znamená dĺžku zvukovej prezentácie jej obsahu. Pri prehľade je táto dĺžka kratšia. To znamená, že užívateľ môže získať rýchly prehľad o obsahu pol minúty pred a pol minúty za aktuálnou pozíciou v priebehu niekoľkých sekúnd. Pri prehľade sa používajú štyri nerečové zvuky. Prvé dva sú úder, jeden slabší, druhý zvýraznený, ktoré informujú o časovej pozícii. Každý piaty úder je zvýraznený. Ďalšie dva reprezentujú odkaz a koniec vety. Takýto prehľad je možné vyvolať kedykoľvek.



Obrázok 7: Rýchly prehľad o obsahu.

Navigácia v tomto programe má štyri režimy. V prvom režime je postupne čítaný celý obsah dokumentu, v druhom si užívateľ prezerá odkazy, v ďalšom nadpisy a v poslednom je čítaný iba textový obsah bez informácie o štruktúre.

5.7.4 Jaws

Jaws je komerčný softvérový produkt pre platformu Windows od spoločnosti Freedom Scientific. Všeobecne je považovaný za najpoužívanejší a predstavuje jedno z najkvalitnejších riešení. Použitím vlastného syntezátora číta informácie z obrazovky. Taktiež je schopný výstupu na aktívne zariadenie v podobe Brailleovho písma. Obsahuje viacero jazykov: angličtinu, španielčinu, francúzštinu, nemčinu, taliančinu, portugalčinu a fínštinu. Dokáže čítať obsah okien programov pre Windows, ktoré používajú štandardné vizuálne rozhranie. S tými najpopulárnejšími pracuje osobitne, aby práca s nimi bola čo najjednoduchšia. Na ovládanie počítača postačuje klávesnica.

S bežnými programami pracuje tak, že po spustení aplikácie prečíta názov hlavného okna. Potom číta názvy okien, ktoré sú potomkami hlavného

okna a zároveň niektorý z ich potomkov obsahuje okno, ktoré zachytáva výstup z klávesnice. Takýmto spôsobom sa užívateľ dozvie, kam píše text, alebo ktorý prvok z programu sa práve klávesnicou ovláda. Jaws informuje o každej stlačenej klávese prečítaním jej znaku, alebo funkcie. Intuitívnym spôsobom číta obsah menu, dialógových okien, a iných bežných prvkov operačného systému. Dôležité je, aby programy obsahovali len štandardné okná poskytované Windows API. S oknami, ktoré patria do neznámej triedy, si Jaws neporadí.

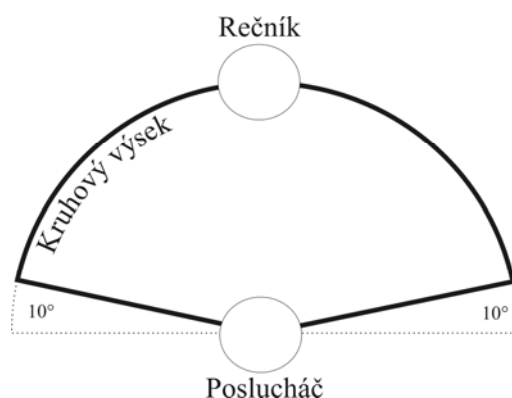
Jedným z programov, pre ktoré je Jaws špeciálne prispôbený je aj Microsoft Internet Explorer. Základným predpokladom pre správne fungovanie Jaws pri čítaní obsahu HTML stránky je jej syntaktická a sémantická správnosť. Čítačka po načítaní stránky prečíta základné informácie o nej - názov a počet odkazov, ktoré obsahuje. Ďalej sa dá stránka spracovať rôznymi spôsobmi. Jaws totižto využíva mnoho klávesových skratiek, ktoré pohyb po stránke uľahčujú. Skákať sa tá medzi odkazmi, pričom je možnosť prechádzať po všetkých, alebo len po už navštívených (kláves V), alebo naopak len po ešte nenavštívených odkazoch (kláves U). Iná navigácia môže byť po všetkých nadpisoch (kláves H), alebo sa dá pohybovať medzi nadpismi konkrétnej úrovne (klávesy 1 až 6). Pre názornosť sme spomenuli len pár zo spôsobov. Ostatné hľadajú na stránke objekty ako sú tabuľky, rámce, zoznamy, obrázky, formuláre a iné.

Celý systém Jaws je komplexný a poskytuje veľa možností, ale zároveň je aj dostatočne intuitívny na to, aby s ním po krátkom čase vedel pracovať každý.

6 Návrh zobrazenia

Najprv prezentujeme základné myšlienky, ktoré budú použité. Tie sa po implementácii predvedú nevidiacim osobám, ktoré ho ohodnotia a navrhnu zlepšenia. Vychádzať budeme z návrhu vytvoreného pre spoločnosť Siemens, ktorý sme popísali v kapitole „A 3D Audio Only Interactive Web Browser„. Jeho myšlienky využívajú všetky prednosti priestorového zvuku, rečové aj nerečové zvuky, riešia základné problémy s odkazmi, poskytujú možnosti zbežného prezretia dokumentu ako aj jeho podrobnú analýzu. Taktiež sa pokúsime pomocou vhodného rozmiestnenia zvukov podporiť proces zapamätávania si do dlhodobej pamäte človeka. Pri získavaní zbežného prehľadu o obsahu využijeme „cocktail party effect“.

Nerečové zvuky sú rozmiestňované, tak ako to je v pôvodnom programe, po obvode zmenšeného polkruhu ležiaceho v horizontálnej rovine. Polkruh sme zmenšili z oboch strán o 10 stupňov. Na hodnotu 10-tich stupňov sme prišli experimentálne. Zvuk sme umiestnili do najľavejšej pozície a pomaly sme ho posúvali po obvode kruhu v smere dopredu. Keď sme subjektívne zaznamenali zmenu pozície, tak sme si aktuálnu polohu zaznamenali. Výsledná hodnota je priemerom niekoľkých meraní. Nejde o významnú veličinu, preto sme spôsob jej získania zbytočne nekomplikovali. Rečové zvuky sú vždy čítané z pozície priamo pred poslucháčom. Svoju polohu nijak nemenia. Dôvodom je subjektívny nepríjemný pocit keď je text čítaný po dlhšiu dobu zľava alebo zprava. Jedine funkcia globálneho prehľadu, ktorej popis je neskôr, umiestňuje rečové zvuky mimo stredovú pozíciu, dokonca aj mimo kruhový výsek.

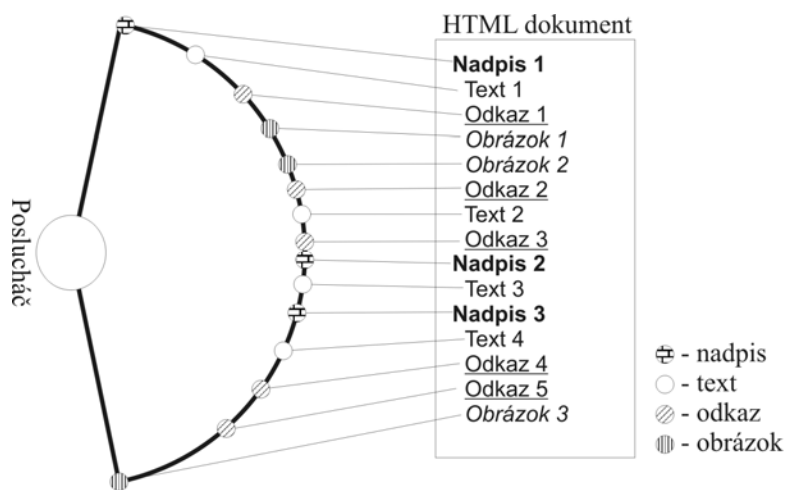


Obrázok 8: Kruhový výsek (pohľad zhora)

Dokument sa pred samotným prezeraním musí spracovať. Analýzou jeho obsahu sa za pomoci syntézy reči vytvoria zvukové objekty, a určí sa ich pozícia v priestore. Objekty sú štyroch typov:

- Nadpis – zodpovedajúci obsahu HTML elementov H1, H2 a H3.
- Odkaz – zodpovedajúci obsahu HTML elementu A.
- Obrázok – zodpovedajúci popisu k obrázku (atribút ALT HTML elementu IMG).
- Text – ľubovoľný iný element.

Program spracúva HTML elementy, ktoré sú viditeľné pri vizuálnom prezeraní obsahu. Výnimku tvoria obrázky, pri ktorých sa ako text použije popis k nim. Informácie o rozdeľovaní textu na jednotlivé objekty a syntéze sú v kapitole o implementácii. Pozícia objektu na kruhovom výseku je úmerná pozícii v texte, ktorá je daná počtom znakov od začiatku dokumentu až po samotný objekt. Prvý objekt je umiestnený na ľavom a posledný na pravom konci výseku. Všetky ostatné sú rozmiestnené medzi nimi.



Obrázok 9: Zobrazenie objektov

Obsah každého objektu tvorí text. Ten je prevedený do zvukovej podoby rečovým syntezátorom. Pre každý typ objektu máme iný syntezátor. Snažili sme sa získať syntezátory s čo najodlišnejšou farbou hlasu pre jasné rozpoznanie typu. Zmysel rôznych syntezátorov je práve v tom, aby užívateľ bez nutnosti vyvolania akejkoľvek inej akcie bol schopný pre daný objekt určiť jeho typ len na základe zvukových vlastností. Transformáciu textu na zvuk sme spravili asynchrónne, znížili sme tak reakčný čas, keďže syntéza je proces výpočtovo náročný. Podrobnejšie sa o syntéze dočítate v kapitole o implementácii.

Okrem hlasu syntezátora, ktorý obsah objektu číta, sa objekty líšia aj v použitých nerečových zvukoch. Každému objektu prislúcha jeden krátky zvuk. Ten svojim charakterom informuje o type a svojou plohou v priestore podáva informáciu o pozícii objektu v dokumente. Narozdiel od rečových, nerečové zvuky nie sú umiestnené v strede, ale majú svoju polohu pozdĺž kruhového výseku. Spôsob umiestnenia bol popísaný na začiatku kapitoly. Keďže aj nerečové zvuky znejú pre rôzne typy objektov odlišne, pomáhajú spolu s farbou hlasu rozoznať o aký typ objektu ide. Tento spôsob zdvojennej identifikácie bol v predhádzajúcich prácach hodnotený pozitívne a zvýšil

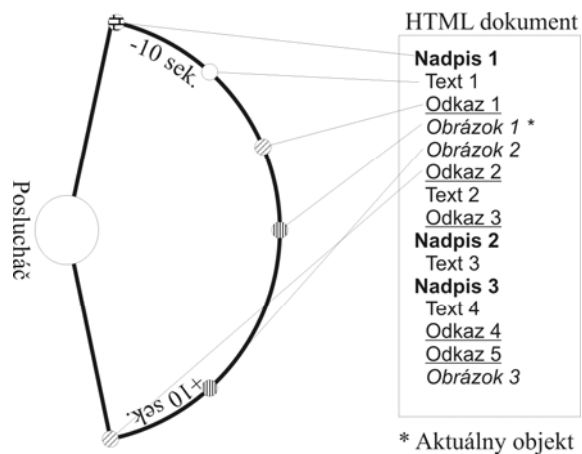
štatistickú úspešnosť správneho rozpoznania obejktu. Najužitočnejším bol nerečový zvuk najmä v prípade keď čítaný text bol krátky a užívateľ nebol schopný za tak krátky čas správne rozoznať farbu hlasu.

Použijeme tri filtre prezerania obsahu. V prvom nie sú žiadne informácie filtrované, a teda je čítaný celý obsah dokumentu. V druhom sú čítané iba odkazy, v treťom nadpisy. Predvolený je prvý z nich. Filtre sú nevidiacimi často používané. Prístup len k nadpisom umožňuje orientačné získanie prehľadu o obsahu. Pre rýchlu navigáciu sa používa filter na odkazy, kedy nie sú čítané zbytočnosti, a je tak urýchlený proces vyhľadávania odkazu.

Najzákladnejšie funkcie sú presun na predchádzajúci objekt, na nasledujúci objekt, funkcia automatického čítania obsahu a zastavenie čítania. Na začiatku po načítaní dokumentu je pozícia nastavená na prvý objekt. Výnikou je len prípad, ak odkaz, ktorým sme sa na dokuemnt dostali, odkazoval na nejaké miesto na nej. Po vyvolaní funkcie na presun na predchádzajúci alebo nasledujúci objekt sa nájde najbližší objekt v danom smere, ktorý vyhovuje podmienke zvoleného filtra obsahu. Ak taký objekt neexistuje, pozícia sa nemení. Automatické čítanie obsahu je iba automatizované posúvanie pozície na nasledujúci objekt hneď po prečítaní obsahu aktuálneho objektu. Taktiež sa uplatňuje použitie filtra. Funkcia zastavenia čítania sa dá vyvolať kedykoľvek a okamžite zastaví zvukový výstup. Je použiteľná aj pri funkciách lokálneho a globálneho prehľadu.

Lokálny prehľad má za cieľ informovať o objektoch nachádzajúcich sa v blízkom okolí aktuálnej pozície. Objekt je v blízkom okolí, ak súčet dĺžok zvukov medzi ním a aktuálnym objektom je menší ako stanovená hranica. Pod dĺžkou zvuku sa rozumie doba čítania v sekundách. Hranicu sme experimentálne stanovili na 10 sekúnd. Pri hľadaní blízkych objektov sa nepoužíva filter obsahu. Keď už sú blízke objekty nachádzajúce sa či už pred alebo za aktuálnym objektom známe, určí sa ich pozícia na kruhovom výseku. Princíp je rovnaký ako pri bežnom prezeraní obsahu. Pri lokálnom prehľade je

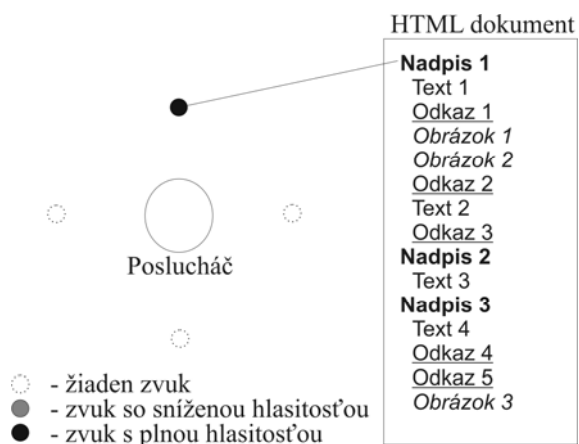
spomínaných 20 sekúnd (10 pred a 10 za aktuálnou pozíciou) transformovaných na celý obvod kruhového výseku. V strede výseku sa nachádza aktuálny objekt. Práve zvýšenie rozlíšenia je zmysel tejto funkcie. Tým, že sa pozície objektov transformujú, získajú od seba väčší rozostup a sú ľahšie identifikovateľné. Navyše pri lokálnom prehľade nie sú čítané obsahy objektov, znejú iba ich príslušné krátke nerečové zvuky. Prehľad začne prvým najľavejším objektom a po pol sekunde je aktivovaný ďalší objekt. Preto celý prehľad, v závislosti na počte blízkych objektov, netrvá viac ako pár sekúnd. Po skončení prehľadu sa objekty nastaví do pôvodných pozícií.



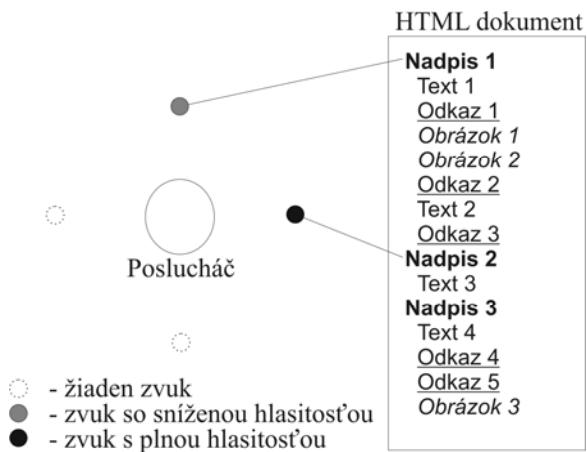
Obrázok 10: Lokálny prehľad

Pridáme možnosť získania rýchleho prehľadu o obsahu pre celý dokument, nielen pre jeho časť. Čítané sú všetky nadpisy počínajúc prvým, avšak v priestore sú umiestnené odlišným spôsobom ako je to pri normálnom prezeraní. Zdroj čítaného zvuku sa môže nachádzať na jednej zo štyroch pozícií: priamo pred, vpravo, vzadu alebo naľavo od poslucháča. Prvý nadpis je prečítaný z pozície priamo vpredu. Po uplynutí 750 milisekúnd zvuk zoslabne na polovičnú úroveň a začne sa čítať nasledujúci nadpis s normálnou hlasitosťou. Tentokrát bude umiestnený vpravo. Podobne bude zvuk po krátkej dobe stíšený a dominantným sa stane ďalší nadpis čítaná zozadu. V tomto

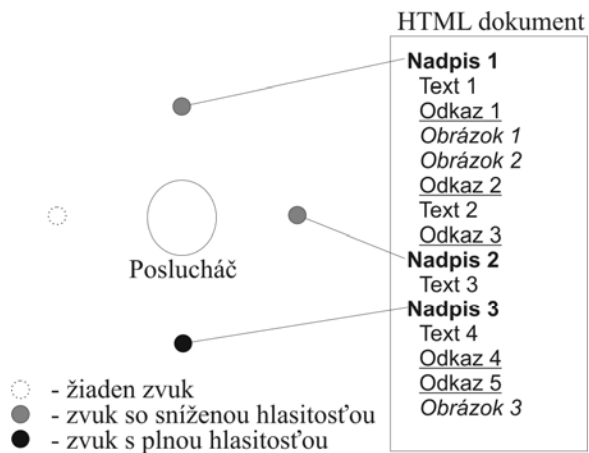
momente sú v priestore aktívne tri súčasne znejúce zvuky (pokiaľ sú dostatočne dlhé), pričom najvýraznejší je ten posledný. Cieľom tohto prístupu je prezentovať obsah dokumentu v čo najkratšom čase. Umiestňovaním zvuku do štyroch vzdialených pozícií a menením intenzity, sa dosiahne dostatočné odlíšenie zvukov, potrebné pre využitie spomínaného “cocktail party effect”-u.



Obrázok 11: Globálny prehľad – prvý nadpis



Obrázok 12: Globálny prehľad – druhý nadpis

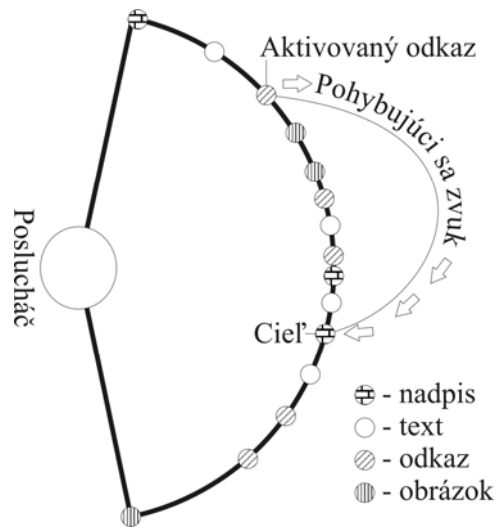


Obrázok 13: Globálny prehľad – tretí nadpis

Nasledovanie odkazov je jednou z najdôležitejších črt HTML prehliadača. My sme zvolili jednoduchý, ale účinný spôsob. Stlačením príslušnej klávesy sa v prípade, že je aktuálnym objektom odkaz, vyvolá navigácia. Tá môže byť dvoch typov v závislosti od typu odkazu:

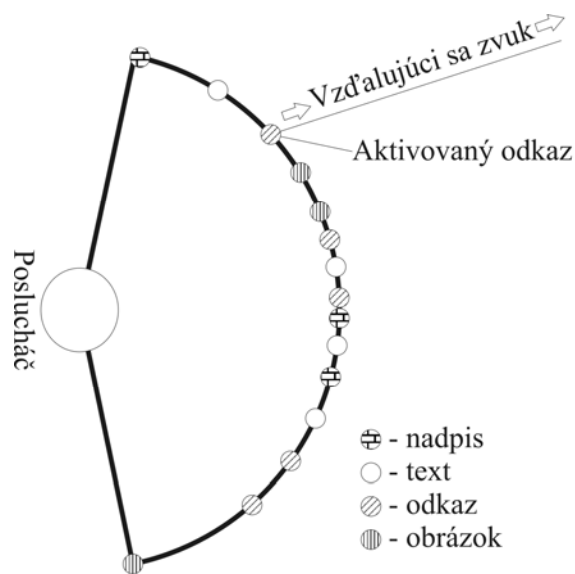
- Vnútoraná
- Vonkajšia

Vnútoraná navigácia je použitá v prípade, že cieľ odkazu sa nachádza v tom istom dokumente ako samotný odkaz. Dokument sa nenačítava odznova, len sa presunie aktuálna pozícia na cieľový objekt. Zvukovo je navigácia sprevádzaná približne dvojsekundovým zvukom, ktorý pripomína vzduchom letiaci predmet. Ten je na začiatku umiestnený na tej istej pozícii na kruhovom výseku ako zdrojový objekt odkazu. Počas dvoch sekúnd sa zvuk rovnomerne presúva na pozíciu cieľového objektu. Podstatné pri navigácii je, aby užívateľ nestratil prehľad o aktuálnej pozícii. Tá je mu týmto spôsobom zreteľne sprostredkovaná. Použitý zvuk má stredovo-výškový charakter a teda je vo zvukovom priestore ľahko lokalizovateľný.

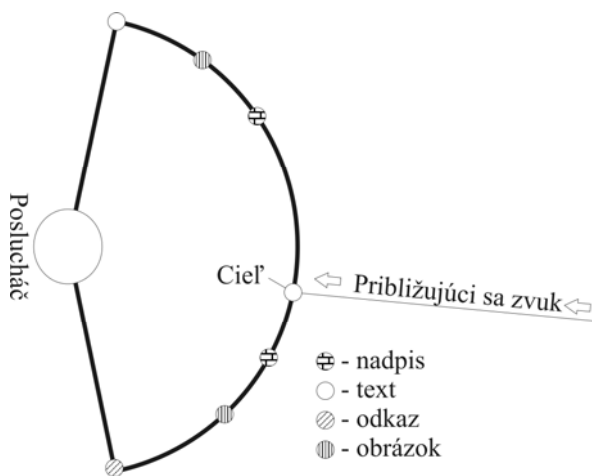


Obrázok 14: Navigácia pre vnútorný odkaz

Vonkajšia navigácia je v prípade, že sa cieľ odkazu nachádza v inom dokumente. Aj v tejto situácii je použitý ten istý zvuk ako pre vnútorný odkaz, avšak ten sa pohybuje po inej trajektórii. Po aktivácii odkazu začne zvuk na pozícii zdroja odkazu na kruhovom výseku a postupne sa vzdáľuje od poslucháča. Vzdáľovanie trvá polovicu času trvania celého zvuku, čiže jednu sekundu. Potom sa začne samotná navigácia. To znamená, že sa načíta nový dokument a identifikujú sa objekty. Keď už je známa pozícia cieľového objektu, začne znieť druhá polovica zvuku a tá sa k cieľovému objektu približuje. Aktuálna pozícia sa nastaví na cieľový objekt.



Obrázok 15: Prvá časť navigácie pre vonkajší odkaz



Obrázok 16: Druhá časť navigácie pre vonkajší odkaz

Spôsob akým sa program ovláda je dôležitou súčasťou návrhu. Aj keď v našom prípade ide o nájdenie vhodného zobrazenia do zvukového priestoru, ovládanie je neoddeliteľnou súčasťou, pretože je úzko prepojené so zobrazením. Aj keď má program štandardné grafické užívateľské rozhranie, a teda ho nevidiaci vedú pomocnými programami ovládať, je nutné aby sa dali

všetky akcie vyvolať pomocou kláves. Iba tak je možné dosiahnuť rýchle prezeranie dokumentu.

Je podstatné, aby správanie programu bolo intuitívne, a aby pohyb zvukových objektov korešpondoval s ovládacími prvkami. Preto základné klávesy, ktorými sú šípky vľavo a vpravo, posúvajú aktuálnu pozíciu smerom vľavo (smer k začiatku dokumentu) alebo vpravo (smer ku koncu dokumentu). Posunutie aktuálnej pozície znamená, že sa aktivuje objekt v danom smere, a teda sa prečíta jeho obsah a zaznie príslušný nerečový zvuk informujúci o pozícii. Podrobnejšie sú klávesové skratky popísané v návode na použitie, ktorý je súčasťou prílohy.

7 Implementácia

Výsledný program je navrhnutý pre platformu Microsoft Windows. Použili sme programovací jazyk C++ ako aj niekoľko voľne šíriteľných programov s prístupným zdrojovým kódom, ktoré implementujú základné metódy ako napríklad: zvukový výstup, rečový syntezátor, k nemu príslušné lexikóny a podobne.

7.1 Festival

Ako sme už spomenuli, ide o voľne šíriteľný syntezátor reči. Pri vývoji vlastného, alebo zlepšenia už existujúceho, spôsobu syntézy reči, bolo treba naimplementovať celý systém. Bola nutná implementácia základnej funkcionality len pre odskúšanie nového nápadu. Práve tento problém bol najdôležitejším dôvodom pre vznik Festivalu. Poskytuje funkcionality pre tri typy užívateľov:

1. Jednoduchá transformácia textu na reč (TTS – anlg. text to speech)
2. Možnosť prídania TTS funkcionality pre programátorov do svojich programov
3. Vývoj nových spôsobov syntézy reči

Viacero spôsobov ovládania:

1. Shell príkazový riadok
2. Programy napísané v skriptovacom jazyku Scheme
3. C++ knižnica
4. Emacs rozhranie

My použijeme tretiu možnosť, a teda Festival pripojíme ako C++ knižnicu. Tým pádom máme nad správaním sa modulu najväčšiu kontrolu, a navyše nie sme príliš ukrátení o funkcionality. Najväčšie možnosti poskytuje Festival cez rozhranie jazyka Scheme.

Festival taktiež využíva iné knižnice. Dve najdôležitejšie z nich sú Speech Tools a interpreter jazyka Scheme - SIOD. Scheme je jazyk podobný Lispu a prostredníctvom programov napísaných v tomto jazyku je možné ovplyvniť všetky fázy procesu syntézy reči. Využíva sa pri vytváraní vlastného lexikónu. Speech Tools tvorí jadro Festivalu a obsahuje základné primitívy nevyhnutné pri syntéze reči.

Festival ako taký je súbor modulov a funkcií pre syntézu reči. Samotný však žiadnu syntézu robiť nevie. Na to je potrebný prídavný modul nazývaný lexikón. Ide o subsystém popisujúci výslovnosť slov. Pozostáva z troch častí:

1. Addenda – krátky zoznam ručne pridaných slov
2. Skompilovaný lexikón – dlhý zoznam (desaťtisíce) slov
3. Metódy popisujúce výslovnosť slov, ktoré nie sú v žiadnom z predchádzajúcich zoznamov

Pre ilustráciu, položky lexikónu majú tvar:

- (“walkers” n (((w oo) 1) ((k @ z) 0)))
- (“lives” n (((l ai v z) 1)))
- (“lives” v (((l i v z) 1)))

Každý lexikón ma svoje jedinečné meno. Festival umožňuje ľubovoľné prepínanie medzi nimi. Líšia sa predovšetkým jazykom, pre ktorý sú vytvorené, a farbou hlasu. V základnej verzii Festivalu sú tri lexikóny pre angličtinu. Každý má inú farbu hlasu a sú rôznej kvality.

Príkaz preloženia slova na text vykoná nasledovné kroky. Najprv sa pozrie do addendy. Ak tam slovo nenájde, hľadá v skompilovanom lexikóne, ak nenájde ani tam, použije pravidlá na preklad slov na reč.

Pravidlá na preklad slov využívajú pravidlá na preklad písmen na reč. Tie majú tvar: (LEFTC [ITEMS] RIGHTC) = NEWITEMS. Príklad: (# [c h] C = k) – ak sa na začiatku slova nachádza „ch“ nasledované spoluhláskou (angl. consonant), tak „ch“ sa vyslovuje ako „k“.

Proces prekladu má nasledovné fázy:

- Rozdelenie vstupu na tokeny – identifikácia začiatku a konca jednotlivých slov, čísel, a pod.
- Identifikácia tokenov – napríklad je vhodné rozlíšiť dátum od bežného čísla, pretože sa čítajú rôznym spôsobom
- Preklad tokenu na slovo (slová) – vhodným príkladom je preloženie číselného zápisu dátumu 18.2.2005 na slová „osemnásteho“, „februára“ a „dvetisícpäť“
- Syntaktická identifikácia slov – podstatné meno sa môže vyslovovať iným spôsobom ako sloveso
- Vyhľadanie v lexikóne
- Intonačný dôraz
- Dĺžka výslovnosti
- Výška tónu
- Vytvorenie výsledného zvuku

Pre naše potreby je dôležité C/C++ rozhranie. To obsahuje nasledovné funkcie:

- `festival_initialize` – inicializácia celého systému, načítanie základných knižníc
- `festival_say_file` – preklad obsahu súboru na zvuk a prehranie na výstupnom zvukovom zariadení
- `festival_say_text` – preklad textu na zvuk a prehranie na výstupnom zvukovom zariadení
- `festival_load_file` – načítanie a spracovanie súboru interpretrom jazyka Scheme
- `festival_eval_command` – vykonanie príkazu jazyka Scheme
- `festival_text_to_wave` – preklad text na zvuk, avšak výstup do pamäťového bloku a nie na výstupné zvukové zariadenie

V našom programe Festival nainicializujeme a pridáme štyri anglické lexikóny. Počas behu, ak si to užívateľ vyžiada, môžeme zmeniť rýchlosť čítania textu. To vykonáme zavolaním funkcie „festival_eval_command“ s príslušným príkazom napísaným v jazyku Scheme. Samotná syntéza je vykonaná volaním funkcie „festival_text_to_wave“.

7.2 OpenAL

Voľne šíriteľné aplikačné rozhranie pre vstupné, ale najmä výstupné, zvukové zariadenie. Umožňuje programátorovi jednoducho vytvárať trojrozmerné scény, simulovať ich akustické vlastnosti, rozmiestňovať zdroje zvukov a dynamicky nimi pohybovať. Poskytuje abstrakciu nad hardvérovou vrstvou a tým pádom odbremeňuje od nutnosti akejkoľvek detekcie zvukového zariadenia. Rozhranie sa podobá na známu grafickú knižnicu OpenGL a vďaka tomu umožňuje jednoduché pridanie zvukov do trojrozmernej grafickej scény.

Z nášho pohľadu je podstatné, že celá simulácia trojrozmerného zvukového priestoru je vykonaná knižnicou. Tá, pokiaľ to výstupné zvukové zariadenie umožňuje, použije optimálnu hardvérovú simuláciu. V opačnom prípade vykoná základné transformácie samotná knižnica. Nie sú také kvalitné, avšak pre naše potreby sú postačujúce. Pod simuláciou sa rozumie zohľadnenie smeru vyžarovania zvuku vzhľadom na poslucháča, útlm hlasitosti s nárastom vzdialenosti, simulácia akustických vlastností hlavy poslucháča, Dopplerov jav, odrazy, obtekanie a ozveny.

Samotné používanie knižnice je jednoduché. Najprv sa celý systém nainicializuje zvolením vhodného výstupného zariadenia (ak je ich viac). Následne sa vytvárajú zdroje zvuku. Zdroj zvuku má niekoľko atribútov, ktoré možno meniť. Má svoju polohu v priestore, smer vyžarovania zvuku, hlasitosť a iné. Samotný zdroj nereprezentuje žiaden svuk. Obsahuje zásobník, ktorý môže obsahovať ľubovoľný počet zvukov. Tie vie postupne prehrať, pričom poradie a priebeh prehrávania je ovládateľný. Preto na prehratie zvuku treba

zvuk najprv vytvoriť, priradiť ho nejakému zdroju zvuku a následne prikázať zdroju, aby daný zvuk prehral. Zvuky vytvárame buď načítaním zo zvukového súboru, alebo ho získame syntézou reči. Samotná OpenAL knižnica však neobsahuje žiadne funkcie na načítanie zvuku zo súboru.

Podobne ako zdroj zvuku, tak aj poslucháč má svoju polohu v priestore. Poslucháč je práve jeden a celá simulácia priestoru je vykonávaná vzhľadom na jeho polohu a smer natočenia.

7.3 ALUT

Knižnica doplnkových funkcií ku OpenAL. Obsahuje funkcie na vytvorenie zvukov priamo zo súborov obsahujúcich nekomprimovaný PCM WAV formát. V našom programe ju používame na načítanie nerečových zvukov.

7.4 Microsoft Internet Explorer

Jediný komerčný produkt, ktorý sme použili je prehliadač internetových stránok Internet Explorer (IE) od spoločnosti Microsoft. Dôvodov je niekoľko. Naš program nemá za úlohu vykonávať lexikálnu, syntaktickú a sémantickú analýzu HTML dokumentu. Predpokladáme, že požadovaný vstup dostaneme od projektu VIPER. Avšak na to, aby sme náš návrh mohli vyskúšať a ohodnotiť, je nutné nejakým spôsobom analýzu dokumentu dosiahnuť. Integrácia tohoto produktu do nášho programu bola najjednoduchšia spomedzi všetkých alternatív a rozhranie, ktoré poskytuje, plne postačovalo našim potrebám. Navyše je IE výborne zdokumentovaný, robustný a poradí si aj s dokumentami, ktoré nie sú syntakticky korektné.

Náš program komunikuje s IE cez COM (angl. Common Object Model) rozhranie. Najprv sa vykoná inicializácia, načítanie HTML dokumentu. O úspešnom spracovaní dokumentu je náš program notifikovaný a začína sa samotné spracovanie obsahu našim programom. Od IE si postupne pýtame

jednotlivé syntaktické elementy, a podľa nich vykonávame potrebné kroky. Či už je to syntéza reči pokiaľ daný element obsahuje text, alebo iný príkaz ovplyvňujúci výslednú podobu transformácie textu na reč.

7.5 B4BLib

Pre potreby zobrazenia HTML dokumentu do zvukového priestoru sme spravili vlastnú knižnicu B4BLib (angl. Browser for Blind Library). Cieľom bolo navrhnúť ju dostatočne všeobecne tak, aby sme mali možnosť jednoducho a efektívne vytvoriť konkrétne zobrazenie. Ide o súbor základných spoločných metód, ktoré sme predpokladali, že výsledné zobrazenie bude potrebovať.

Základnou triedou je trieda IB4BLib, ktorá zabezpečuje inicializáciu knižnice a vytvorenie inštancie triedy IB4BTSMng. Tá má za úlohu spravovať syntezátory reči. Pod správou rozumieme vytvorenie, ovládanie, volanie metód a nakoniec ich uvoľnenie. V súčasnosti B4BLib podporuje iba Festival, avšak vďaka použitej abstrakcii sa dá rozšíriť na akýkoľvek iný. Napríklad o rozhranie Speech API, vyvinuté spoločnosťou Microsoft a dodávané k operačnému systému Windows.

Pri práci so zvukmi používame potomkov triedy IB4BSoundObj, ktorá reprezentuje ľubovoľný zvukový objekt. Či už ide o syntezovanú reč (trieda IB4BSoundObjText), alebo nerečový zvuk (trieda IB4BSoundObjEarcon).

Za veľmi dôležitú vlastnosť celej knižnice považujeme jej asynchrónne vytváranie zvukových objektov. Dôvodom je pomerne pomalá syntéza reči. Užívateľ nesmie byť citelne obmedzovaný pri prezeraní dokumentu zbytočným čakaním na transformáciu celého obsahu na reč. Preto sa všetky zvuky syntezujú na pozadí v samostatnom vlákne. To vyžaduje zložitejší prístup k implementáciám všetkých objektov, zabezpečenie korektnej práce so zdieľanou pamäťou a riešenie bežných problémov pri synchronizácií. Za týmto účelom sme vytvorili triedu IB4BAsyncLoader, ktorá zabezpečuje asynchrónne vytváranie objektov. Tie sú potom uložené v jednej zo stránok záložnej pamäti

reprezentovanej triedou `IB4BCachePage`. Inštancii tejto triedy môže byť viac a všetky sú združené a prístupné cez rozhranie triedy záložnej pamäte `IB4BCache`. Myšlienka viacerých stránok spočíva v tom, že každý dokument má syntezovanú reč uloženú vo svojej vlastnej stránke. Tá sa dá po skončení prezerania ľahko uvoľniť, pričom ak sú otvorené iné dokumenty, tie v záložnej pamäti zostanú. Špeciálna stránka existuje pre nerečové zvuky, ktoré sa využívajú počas celého behu programu a z pamäte sa uvoľňujú až pri skončení behu programu.

7.6 *B4BReader1Lib*

Knižnica, o ktorej sme hovorili v predchádzajúcej kapitole, poskytovala základné metódy potrebné na implementáciu nášho programu. `B4BReader1Lib` je knižnica, ktorá už predstavuje samotnú implementáciu nášho návrhu zobrazenia obsahu HTML dokumentu do zvukového priestoru. Implementuje rozhranie `IB4BReader` definované v knižnici `B4BLib`.

Najzaujímavejšou triedou je trieda `CPage`, ktorú tvoria základné algoritmy. V prvom rade ide o analýzu obsahu HTML dokumentu pri jeho otvorení. Ako sme už spomínali, na lexikálnu, syntaktickú a sémantickú analýzu používame Microsoft Internet Explorer. Ten nám poskytuje objektový dátový model celého dokumentu. Každý HTML dokument sa dá reprezentovať stromom, ktorého vrcholy sú jednotlivé HTML elementy alebo ich textový obsah. Vrcholy sú usporiadané v tom istom poradí, v akom sa v dokumente vyskútujú. To znamená, že ľavý potomok daného uzla sa vyskútuje v dokumente skôr, ako iný súrodenec napravo od neho. Vďaka tejto reprezentácii prechádzaním stromu algoritmom preorder analyzujeme dokument zhora nadol.

Pred samotnou analýzou ešte prebehne odstránenie nežiadúcich komentárov, ktoré na obsah nemajú žiaden vplyv. Následne sa začne rekurzívne prechádzať strom od koreňa, ktorý je tvorený HTML elementom „Body“.

Keďže pri čítaní obsahu dokumentu rozlišujeme odkazy, nadpisy, obrázky a ostatné texty, je potrebné aby sme v každom momente vedeli, či je nájdený text potomkom jedného z HTML elementov „a“ (odkaz), „h1, h2, h3“ (nadpis) alebo „img“ (obrázok). Ak nejaký text spĺňa túto požiadavku, je celý odoslaný na asynchrónne načítanie knižnicou B4BLib príslušným syntezátorom, keďže pre odkaz, nadpis a obrázok máme rôzne syntezátory.

Ostatné texty sú čítané štvrtým syntezátorom, avšak navyše sú delené na menšie časti. Väčšinou je oddeľovačom znak ukončenia vety. Okrem neho sú to elementy ako „br“, „td“, „tr“, „p“ a taktiež je obmedzená maximálna dĺžka textu vhodne stanovenou konštantou. Dôležité je aby texty na jednej strane neboli príliš dlhé a umožňovali tak dostatočne podrobné prechádzanie obsahu, ale aby zase nebol príliš krátky, čo by znemožnilo syntezátoru reči určiť správne intonáciu pri čítaní. Je ideálne ak syntezátor (v našom prípade Festival) dostane na vstup celú vetu spolu s interpunkčnými znamienkami, vtedy je syntéza najprirodzenejšia.

Globálny prehľad je implementovaný v nezávislom vlákne tak, aby bol vykonávaný nezávisle od užívateľského rozhrania. Ak bolo už spomenuté v popise nášho návrhu, globálny prehľad využíva „cocktail party effect“ na to, aby v čo najkratšom čase prečítal všetky nadpisy na stránke. Podobne ako globálny prehľad bežia v nezávislých vláknach aj lokálny prehľad, nasledovanie odkazu a čítanie celého obsahu dokumentu.

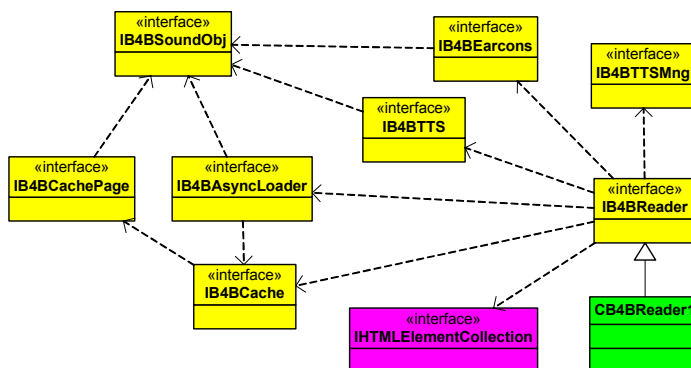
7.7 B4B

Aby sme mohli všetky spomenuté knižnice spustiť a ovládať v prostredí operačného systému Microsoft Windows, je nutné ich spojiť s grafickým užívateľským rozhraním. Aj keď to nie je vyslovene nutné, keďže náš program je v prvom rade pre nevidiacich a jeho idea je práve v tom, že prezerat' obsah HTML dokumentu sa dá aj bez vizuálneho vnemu. B4B obsahuje jedno základné okno, v ktorom je integrovaný prehliadač HTML dokumentov

Microsoft Internet Explorer. Obsahuje menu so všetkými akciami, ktoré sa dajú vykonať aj pomocou stlačenia príslušného tlačidla na klávesnici.

Súčasťou celkového riešenia programu je aj inštalačný balík, ktorý celý program spolu s Festivalom, lexikónmi a OpenAL knižnicou nainštaluje na cieľový počítač. Program sa tak dá jednoducho odstrániť štandardným spôsobom.

7.8 Architektúra



Obrázok 17: Základné vzťahy medzi rozhraniami

8 Testovanie

Po ukončení implementácie bolo nutné naše nápady odskúšať a zistiť ich nedostatky. Už aj počas návrhu a implementácie sme boli s nevidiacimi v kontakte a získavali sme od nich informácie o tom, ako vnímajú priestor, zvuky, ako sa orientujú v priestore a podobne. Zaujímavou skúsenosťou bolo pozorovať ich pri práci s počítačom. Dokážu ho veľmi rýchlo ovládať. Ako sa neskôr počas testovania ukázalo, je to najmä vďaka syntezátoru, na ktorý sú zvyknutí. Ten číta text tak rýchlo, až je pre netrénovaného poslucháča úplne nezrozumiteľný. Ďalšou výkonnou pomôckou sú klávesové skratky. Nimi vedia ľahko vykonať základne operácie nutné pre prácu s programami.

So spoluprácou pri testovaní súhlasili dvaja nevidiaci, jeden slabozraký a šiesti vidiaci. Iba jeden nevidiaci mal krátku skúsenosť s našim program. Ostatní s ním nikdy nepracovali. Obaja nevidiaci používajú pri prehliadaní HTML dokumentov program Jaws. Slabozraký tento program nepoužíva, keďže jemu postačuje zväčšovacie sklo a pomocný program, ktorý zväčšuje objekty na obrazovke. Všetci dostali rovnaké úlohy avšak ku zrakovo postihnutým bol iný prístup. Zaujímali nás nielen objektívne merané veličiny, ale aj ich subjektívny pocit z programu.

Dokumenty boli získané z internetovej encyklopédie Wikipedia. Boli nájdené dokumenty na náhodne zvolené témy. Stránky Wikipédie spĺňajú štandardy a všetky potrebné informácie týkajúce sa danej témy sa nachádzajú v jednom dokumente. Pre potreby testovania boli zo stránok odstránené nadbytočné časti, ktoré prekážali.

Oslovené osoby dostali postupne tri úlohy. Prvá bola určená len pre zrakovo postihnutých. Jej úlohou bolo overiť použiteľnosť funkcie globálneho prehľadu. Druhej a tretej úlohy sa už zúčastnili všetci. Tie boli zamerané na presnosť pri určovaní polohy a na schopnosť rozoznávať hlasy použitých syntezátorov.

Mali sme v úmysle porovnať náš program s najpoužívanejším programom pre nevidiacich Jaws, avšak ako nám nevidiaci povedali, náš program by mal veľký hendikep, pretože oni sú na Jaws a jeho syntezátor zvyknutí. Používajú ho tak, že syntezátor číta veľkou rýchlosťou a majú v ňom nastavené klávesové skratky. Taktiež majú isté zaužívané postupy, ktoré pri prezeraní dokumentu používajú. Objektívne porovnanie by bolo možné len po dlhodobom používaní nášho programu, čo však je problém, keďže dokonca i s programom Jaws vie plnohodnotne pracovať málokto z nevidiacich. Aj napriek tomu bolo zaujímavým výsledkom pri prvej úlohe zistenie, že podľa jedného z nevidiacich by bol náš program v testovanej funkcii lepší.

Použili sme prenosný počítač HP nx6310 vybavený bežnou zvukovou kartou. Tá neposkytovala žiaden nadštandardný priestorový vnem. Podobne boli pri testoch s nevidiacimi použité obyčajné slúchadlá. Vidiace osoby mali k dispozícii kvalitné štúdiové slúchadlá AKG k240, avšak ako sa ukázalo, nijako to výsledky neovplyvnilo.

8.1 Úloha č. 1

Cieľom bolo zistiť, do akej miery sú schopní zachytiť čítaný text pri použití globálneho prehľadu. Ako bolo spomenuté, táto funkcia využíva prekrývajúce sa zvuky aby dosiahla rýchlejšie získanie informácií. Účastník dostal stránku, na ktorej sa nachádzalo niekoľko nadpisov. Vyvolal funkciu globálneho prehľadu a snažil sa zachytiť čo najviac z nich. Podstatné bolo, aby získal prehľad o obsahu stránky a nie aby si nadpisy zapamätal. Táto úloha nebola objektívne merateľná. Pôvodne sme chceli, aby nám zrakovo postihnutí povedali čo najviac nadpisov, ktoré počuli, avšak problémy boli spôsobené skôr pamäťou ako porozumením. Rozhodli sme sa pre číselné hodnotenie od 1 do 5, kde 1 znamená, že obsah stránky je veľmi málo jasný a 5 bodov bolo udelených stránke, ktorej obsah bol kompletne pochopený. Stránok bolo celkovo 10, a teda bolo pri tejto úlohe vykonaných 30 meraní.

Osoba	1	2	3	4	5	6	7	8	9	10
1	2	4	4	4	5	3	5	5	5	5
2	4	4	5	5	5	5	5	2	5	5
3	5	5	4	5	5	5	5	3	5	5

Tabuľka 1: Výsledky úlohy č. 1

Priemer	
1	4,2
2	4,5
3	4,7
Celkový priemer	4,5

Tabuľka 2: Priemerné hodnotenie pre osoby

Výsledky sú pozitívne. Zrakovo postihnutí nemali takmer žiadne problémy pri porozumení, aj keď sa čítané texty prekrývali. Ak sa problémy vyskytli, tak boli spôsobené najmä nedostatkami v znalosti anglického jazyka.

Prvá osoba uviedla, že je pre ňu máťúce keď sa niekoľko za sebou idúcich nadpisov začína na rovnaké slová. Vtedy sú podstatné práve koncové časti nadpisov, avšak tie sa stratia v zhluku ostatných slov. Podľa jeho názoru by mu funkcia globálneho určite zrýchlila získanie prehľadu o obsahu stránky v porovnaní s programom Jaws. Je dôležité podotknúť, že dotyčná osoba vie s programom Jaws pracovať najzručnejšie spomedzi oslovených.

Druhý nevidiaci nevie, či by funkciu rýchleho prehľadu využil, pretože získanie rýchleho prehľadu nepotrebuje. Podľa jeho slov si stránku postupne prezrie od začiatku po koniec, a to mu netrvá tak dlho, aby na to potreboval pomoc.

Posledný z účastníkov prvej úlohy mal najmenšie problémy pri porozumení čítaného textu. Ako rušivé vnímal to, keď sa sústredil na čítaný nadpis a začal sa čítať ďalší. Často krát to spôsobilo odpútanie pozornosti, avšak výsledkom nebolo nepochopenie významu. Tento problém nás zaujal, a tak sme ho skúsili odstániť tým, že sa hlasitosť čítania nebude meniť. Ako

bolo pri popise globálneho prehľadu spomenuté, posledný čítaný nadpis má raz tak vysokú úroveň hlasitosti ako ostatné. Toto sme odstránili a hlasitosti sme nemenili. Výsledok sme dali účastníkovi vyskúšať a hodnotil ho pozitívne. Bol schopný dopytovať čítaný text do konca bez toho, aby bol vo veľkej miere rušený nasledujúcim nadpisom. Keď nás zaujímalo, či by použitie viac odlišných hlasov pomohlo, jeho názor bol, že by to naopak situáciu zhoršilo, keďže by na seba príliš pútali pozornosť. Je to síce v rozpore s predpokladmi použitia „cocktail party effect“-u kde je dobré ak sú hlasy rôzne, avšak súhlasíme s tým, že všetky súčasne znejúce hlasy by mali byť v istom zmysle uniformné. Nesmie byť medzi nimi žiaden dominantný, ktorý by svojim zvukom prekryval ostatné. To je však ťažko objektivizovateľná veličina.

8.2 Úloha č. 2

Tejto úlohy sa zúčastnilo všetkých deväť oslovených osôb. Zaujímalo nás ako presne sú schopní určiť polohu zvuku v priestore. Osoba mala na stránke nájsť vopred známy nadpis a povedať v percentách kde sa na stránke nachádza. Použili sme tie isté dokumenty ako v prvej úlohe, avšak rozšírili sme ich počet na 15. Na každej stránke sme náhodne vybrali jeden nadpis a preň sme si zistili správnu pozíciu. Nadpisy aj stránky boli pre všetky osoby rovnaké. Nadpis úplne na začiatku je na pozícii 0% a posledný nadpis má pozíciu 100%. Pri tejto úlohe sa nepoužívala žiadna špeciálna funkcia programu, iba sa pomocou základných ovládacích prvkov našiel hľadaný nadpis. Ako bolo spomenuté pri popise systému, pozícia objektu je reprezentovaná pozíciou nerečového zvuku na kruhovom výseku.

Osoba	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	12	60	35	15	65	50	25	50	50	30	0	40	20	15	75
2	30	60	35	15	80	50	25	65	60	30	0	50	15	25	65
3	37	60	35	10	85	50	25	45	45	30	0	45	20	25	60
4	25	55	25	15	70	55	25	50	55	20	0	48	10	10	70
5	30	80	30	10	80	60	30	50	70	30	0	50	10	10	60
6	15	55	30	5	90	50	20	85	55	40	0	55	5	35	60
7	15	60	10	5	80	40	30	50	50	10	0	40	5	10	60
8	15	65	35	5	75	50	25	50	60	40	0	55	5	35	75
9	30	60	30	15	65	45	20	75	50	40	0	50	15	25	65

Tabuľka 3: Namerané hodnoty (v %)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
26	65	30	6	81	54	14	61	61	28	0	48	17	15	64

Tabuľka 4: Skutočné hodnoty (v %)

0	0%
6	5,2%
14	11%
15	9,4%
17	6,7%
26	7,9%
28	7,8%
30	5%
48	4,3%
54	5,6%
61	13%
61	8%
64	5,1%
65	6,7%
81	7,2%

Tabuľka 5: Odchýlky pre pozície

Osoba	
1	7,3%
2	4,1%
3	6,5%
4	6,1%
5	5,7%
6	8,4%
7	9,1%
8	7,5%
9	6,7%
Celkový priemer	7,9%

Tabuľka 6: Odchýlky pre osoby

Z výsledkov vyplýva, že presnosť určovania pozície je pomerne dobrá. Keďže celkový priemer odchýlok je 7,9%, dá sa z toho usúdiť, že na kruhovom výseku sa dá spoľahlivo rozlíšiť $100/7,9 = 12,5$ pozícií. To je v súlade s výsledkami predchádzajúcich prác, ktoré uvádzali mierne horšiu rozlišovaciu schopnosť na úrovni okolo 10-tich pozícií. V tabuľke odchýlok pre pozície sa dá pozorovať rovnomerné rozdelenie. Očakávali sme menšie nepresnosti v oblasti priamo pred poslucháčom. Tam má človek najväčšiu schopnosť presne zamerať pozíciu.

Použitý nerečový zvuk signalizujúci pozíciu bol pre zrakovo postihnutých príjemný a zreteľný. Neprekážalo im, že ho počujú pri každom zvukovom objekte. Všetky osoby okrem jednej vnímali zvuk tak, že sa pohybuje v horizontálnej rovine pred nimi. Iba jedna vidiaca osoba mala pocit, že je zvuk za ňou, čo jej sťažovalo odhadovanie polohy, ale v štatistike sa to negatívne neprejavilo. Aj to je dôkaz spomínaných pred-zadných nejasností a psychologického vnemu, že ak človek niečo počuje ale nevidí, tak sa to zrejme nachádza za ním.

8.3 Úloha č. 3

Zaujímalo nás taktiež, či sú hlasy použitých syntezátorov dostatočne odlišné. Pri tejto úlohe bolo zúčastnených osem osôb. Každá mala pred samotnou úlohou k dispozícii dokument, na ktorom sa snažili naučiť, ktorý hlas prislúcha ktorému typu objektu. Potom dostali dokument s náhodnými objektami a mali podľa hlasu povedať, akého typu sú. Objektov bolo v dokumente 30.

Osoba	
1	100,00%
2	86,67%
3	56,67%
4	86,67%
5	46,67%
6	50,00%
7	60,00%
8	46,67%
Celkový priemer	66,67%

Tabuľka 7: Úspešnosť pre osoby

Typ objektu	
Nadpis	68,30%
Odkaz	55,10%
Obrázok	63,10%
Text	68,85%

Tabuľka 8: Úspešnosť pre typy objektov

Pri tejto úlohe nás výsledky neuspokojili. Celková úspešnosť na úrovni 67% je pomerne nízka. Pozitívne však pôsobí fakt, že prvá osoba nezaznamenala ani jeden nesprávny úsudok. Ide o nevidiaceho, ktorý si program vyskúšal už skôr ako ostatní. V súlade s týmto pozorovaním je skutočnosť, že ostatní účastníci sa počas testovania sťažovali na fakt, že si jednoducho napamätajú ktorý hlas prislúcha ktorému objektu. Je teda

predpoklad, že po dlhšej práci s programom by sa úspešnosť významne zvýšila. Ťažkosti boli najmä s krátkymi slovami, alebo so slovami, s ktorými mali syntezátory problémy pri výslovnosti.

Ideu použitia viacerých hlasov hodnotili zrakovo postihnutí pozitívne. Ľahko tak získajú informáciu navyše.

8.4 Výsledky

Celkovo z testov vyplynula použiteľnosť našich ideí. Zlepšenie by malo nastať najmä v smere použitých syntezátorov. Ich kvalita bola síce na vysokej úrovni, keďže účastníci nemali pri prvej úlohe problém im rozumieť, avšak neboli vzájomne dostatočne odlišné. Výroba syntezátora je ale ťažká úloha a je mimo rozsah našej práce.

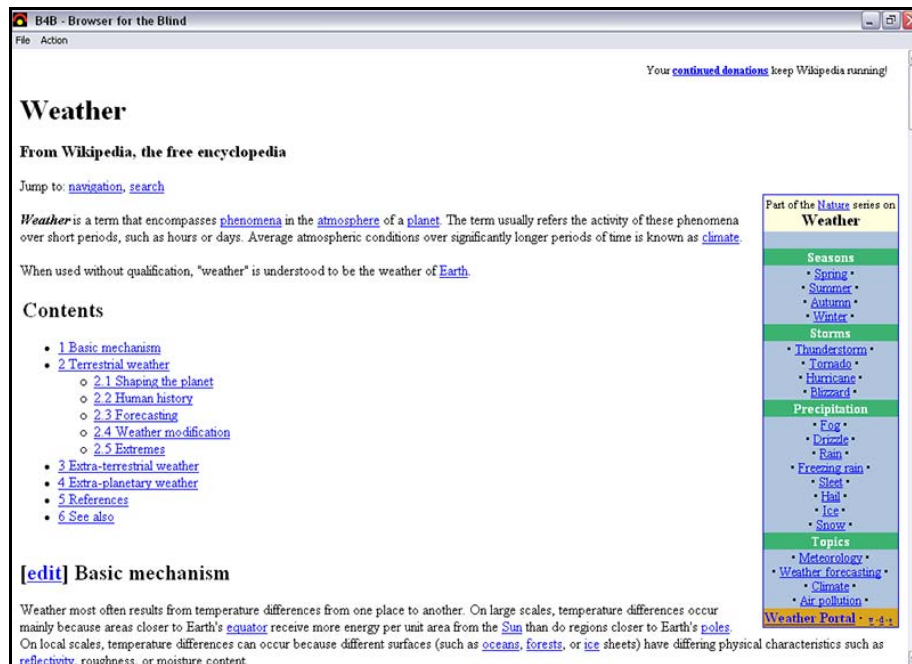
9 Prílohy

9.1 Návod na použitie

9.1.1 Inštalácia

Spustením inštalačného súboru sa program a jeho súčasti skopírujú na cieľový počítač a vytvoria sa odkazy pre jeho štandardné spustenie. Taktiež je možné program jednoducho z počítača odstrániť. Po úspešnom skončení inštalačného procesu, je možné program spustiť cez odkaz v menu „Štart“.

9.1.2 Uživateľské rozhranie



Obrázok 18: Ukážka programu

Program obsahuje hlavné okno, ktoré pozostáva z integrovaného prehliadača Microsoft Internet Explorer. V ňom sa zobrazuje obsah prezeraného

dokumentu. Ovládanie zabezpečuje menu spolu s klávesovými skratkami, ktorých použitie nevidiaci uprednostňujú. Nasleduje popis menu:

- File
 - About – Základné informácie o programe
 - Exit (x) – Ukončenie práce s programom
- Action
 - Go to (g) – Zobrazí dialógové okno, do ktorého užívateľ zadá URL dokumentu, ktorý si želá zobrazit'
 - Follow link (enter) – Spustenie navigácia aktívneho odkazu
 - Back (backslash) – Navigácia na predchádzajúci dokument
 - Forward (alt+backslash) – Navigácia na nasledujúci dokument
 - Refresh (F5) – Obnovenie dokumentu
 - Slower (-) – Spomaľenie čítania
 - Normal (*) – Obnovenie prednastavenej rýchlosti čítania
 - Faster (+) – Zrýchlenie čítania
 - Preview (p) – Lokálny prehľad
 - Global preview (o) – Globálny prehľad
 - All (a) – Režim čítania celého obsahu
 - Links (l) – Režim čítania iba odkazov
 - Headings (h) – Režim čítania iba nadpisov
 - Next (right) – Nasledujúci objekt
 - Previous (left) – Predchádzajúci objekt
 - Page up (page up) – Preskočenie niekoľkých objektov v smere k začiatku dokumentu

- Page down (page down) – Preskočenie niekoľkých objektov v smere ku koncu dokumentu
- Begin (home) – Skok na začiatok dokumentu
- End (end) – Skok na koniec dokumentu
- Start (space) – Automatické čítanie obsahu
- Stop (escape) – Ukončenie čítania

10 Reference

- [1] Brungart, D. S. and Simpson, B. D., Optimizing a Virtual Speech Display: Comments on Brungart and Simpson, ICAD 2003, ACM, 2005
- [2] Brungart, D. S. and Simpson, B. D., Optimizing the Spatial Configuration of a Seven-Talker Speech Display, ACM, 2005
- [3] Talker Speech Display, ACM, 2005
- [4] Evans, M. J., Synthesizing Moving Sounds, ICAD, 1998
- [5] Festival, <http://www.cstr.ed.ac.uk/projects/festival/>
- [6] Gardner, B. and Martin, K., HRTF Measurements of a KEMAR Dummy-Head Microphone, MIT Technical Report #280, 1994
- [7] Goose, S. and Möller, C., A 3D Audio Only Interactive Web Browser: Using Spatialization to Convey Hypermedia Document Structure, ACM, 1999
- [8] Cheng, C. I. and Wakefield, G. H., Moving Sound Source Synthesis for Binaural Electroacoustic Music Interpolated Head-Related Transfer Functions (HRTFs), Computer Music Journal, 2001
- [9] Jaws, <http://www.freedomscientific.com/>
- [10] Kobayashi, M. and Schmandt, Ch., Dynamic Soundscape: mapping time to space for audio browsing, ACM, 1997
- [11] Lokki, T., Gröhn, M., Savioja, L. and Takala, T., A Case Study of Auditory Navigation in Virtual Acoustic Environments, Helsinki University of Technology
- [12] Lumbreras, M. and Rossi, G., A Metaphor for the Visually Impaired: Browsing Information in a 3D Auditory Environment, ACM, 1995
- [13] Neuhoff, J. G., Kramer, G. and Wayand, J., Sonification and the Interaction of Perceptual Dimensions: Can the Data get Lost in the Map?, ICAD, 2000

[14] Shinn-Cunningham, B., Learning Reverberation: Considerations for Spatial Auditory Displays, ICAD, 2000

[15] Schmandt, Ch., Audio Hallway: a Virtual Acoustic Environment for Browsing, ACM, 1998

[16] Schmandt, Ch., and Mullins, A., AudioStreamer: Exploiting Simultaneity for Listening, ACM, 1995

[17] Wenzel, E. M., Wightman, F. L. and Kistler, D. J., Localization with non-individualized virtual acoustic display cues, ACM, 1991

[18] Wikipedia, <http://www.wikipedia.org/>