

Ing. Ivan Kišac

Animácia ľudskej tváre a vizuálna reč

(Speaking head)

Vedúci diplomovej práce: Ing. Peter Drahoš

Obsah

- Úloha
- Využitie
- Existujúce riešenia
- Analyzované oblasti
- Návrh spracovania
- Spracovanie textu
- Produkcia výstupu
- Príprava kľúčových snímok
- Tok údajov
- Implementácia
- Výsledky práce
- Ďalšie možné vylepšenia

Úloha

- Analyzujte možnosti a techniky tvorby vizuálnej reči v reálnom čase, zamerajte sa na analýzu konverzačných systémov využívajúcich 3D avatarov na reprezentáciu používateľov.
- Implementujte aplikáciu schopnú generovať a vizualizovať reč z textového či audio zdroja v reálnom čase s využitím existujúcich text-to-speech systémov a vizualizačnej knižnice ľudskej tváre.

Využitie animácie ľudskej tváre

- Zábavný priemysel
 - filmy
 - hry – náročné na výkon
- Používateľské rozhranie
 - webové používateľské rozhranie
- Pedagogika
 - výučbové programy
 - cudzie jazyky
 - výučba nepočujúcich
- Medicína



Existujúce riešenia

- proFace
 - animačný nástroj
 - parametrické deformácie
 - animácia v reálnom čase
- Crazytalk
 - parametrické deformácie z obrázku
- Baldi a Timo
 - „struny“
 - výučbový softvér

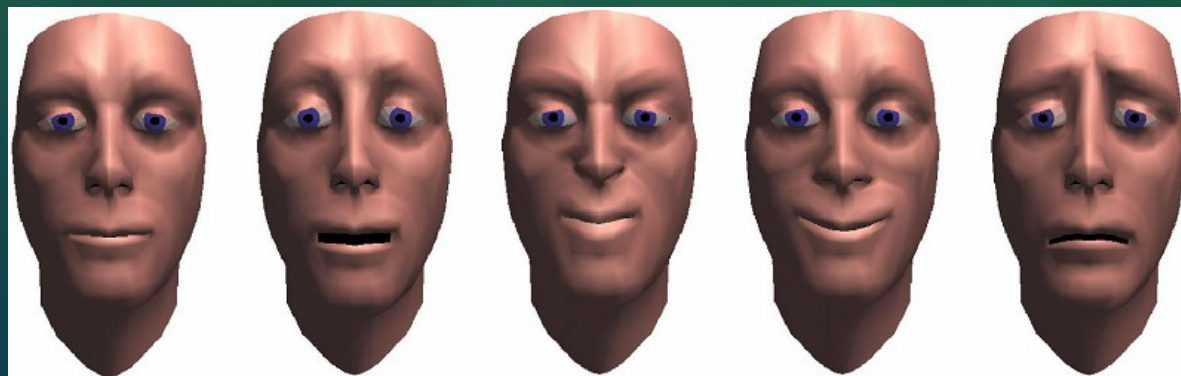


Ciele práce

- Spracovanie vstupného textu a animácia modelu na jeho základe
- Zakomponovanie
 - emócií
 - znakov ľudskej mimiky
- Použitie vizualizačnej knižnice LibFace
- *Zvukový výstup reči*
 - *synchronizácia*
- *Rozšíriteľnosť*

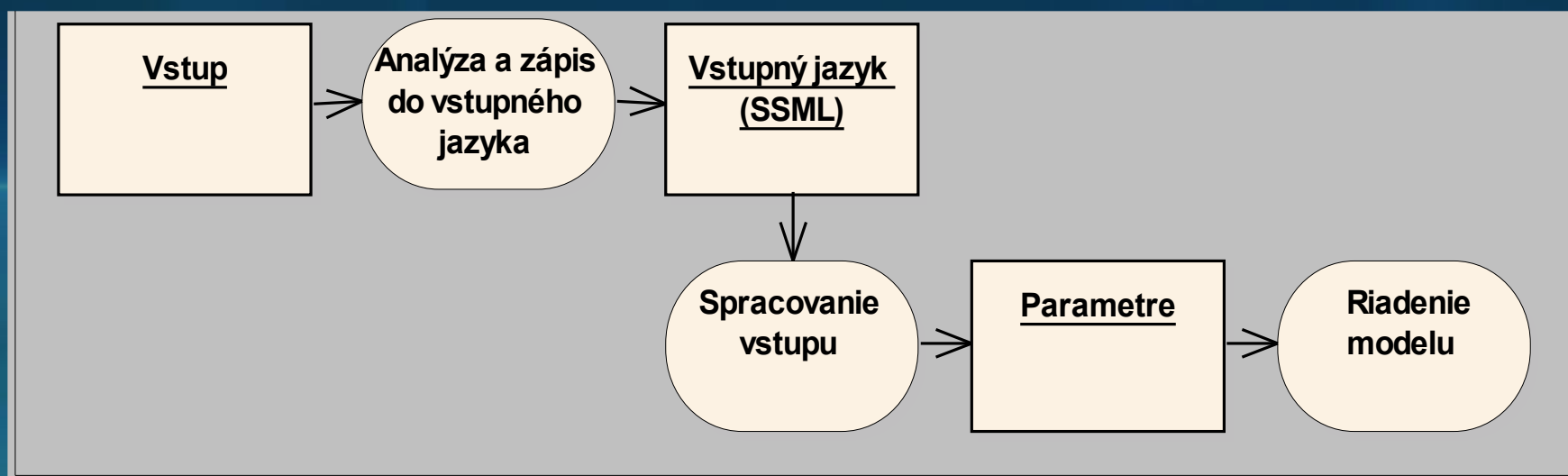
Analyzované oblasti

- animácia
- psychológia
 - emócie / nálady
 - mimojazykové rečové prejavy
 - pohyby hlavy
 - pohyby očí, mrkanie, ...
- fonetika
 - stavebné prvky reči
 - fonémy, difóny a vizémy
 - intonácia



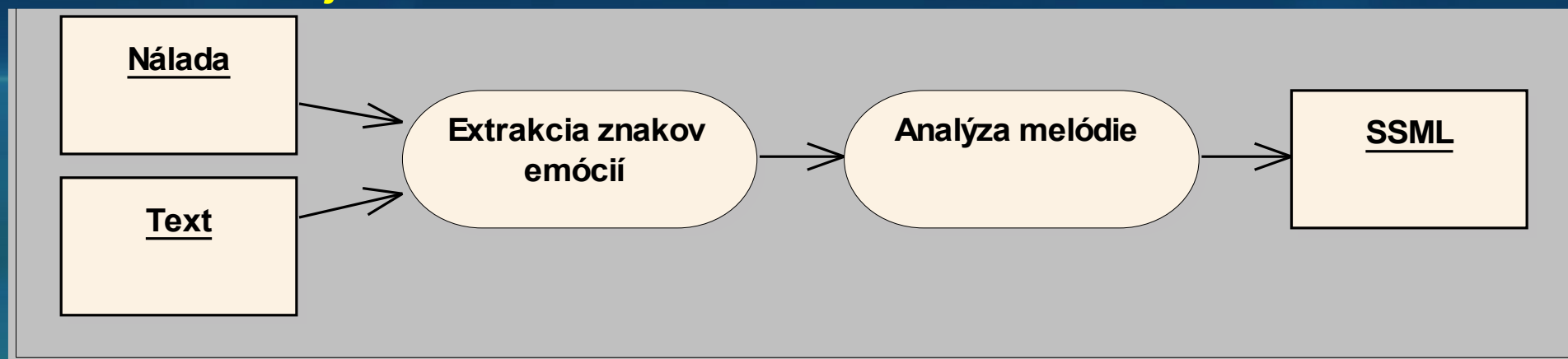
Návrh spracovania

- architektúra – dátovody a filtre
 - rôzne vstupy
 - transformácia na jednotný vstupný jazyk SSML
 - spracovanie do parametrov pre riadenie modelu



Spracovanie textu

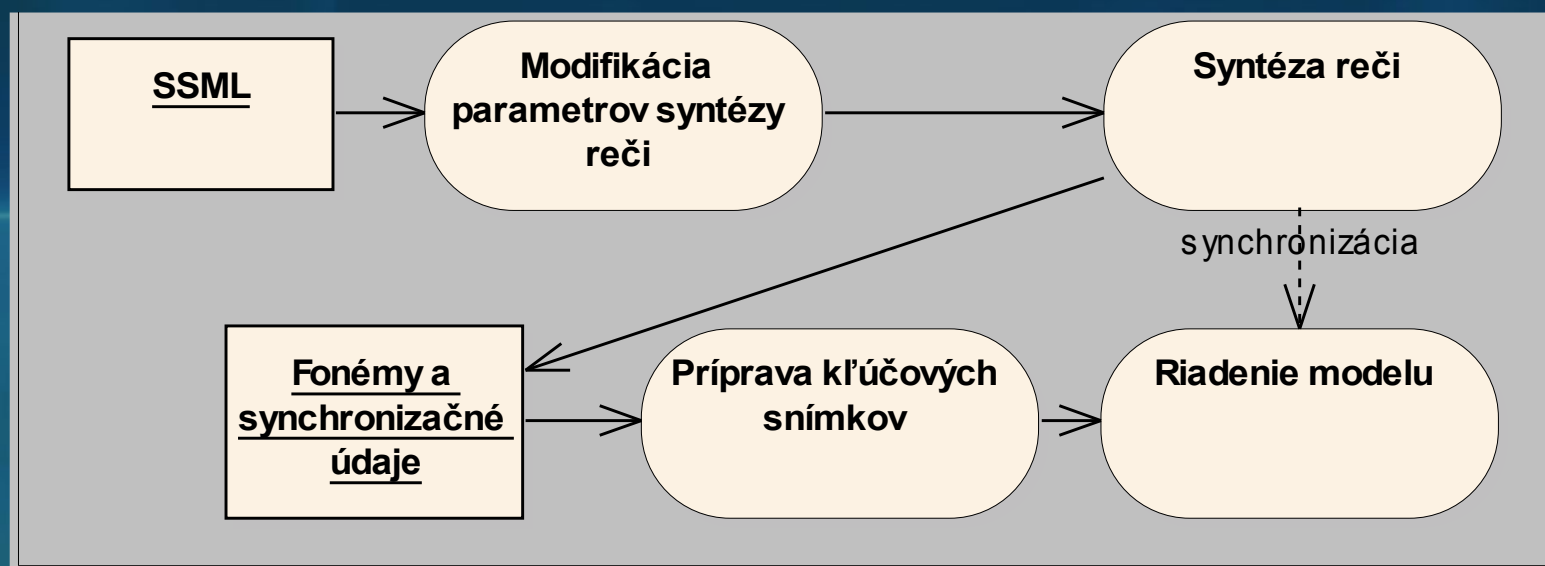
- Odstránenie značiek emócií (smajlíkov)
 - nahradenie SSML značkami
 - :) → <mark name="smile"/>
- Analýza melódie
 - doplnenie SSML značiek označujúcich interpunkciu
 - Ahoj! <mark name="exclamation"/>



Produkcia výstupu

- eSpeak

- analýza SSML
- modifikácia parametrov syntézy reči
- syntéza reči → audio údaje, udalosti (časové umiestnenie v údajoch, názvy foném, ...)

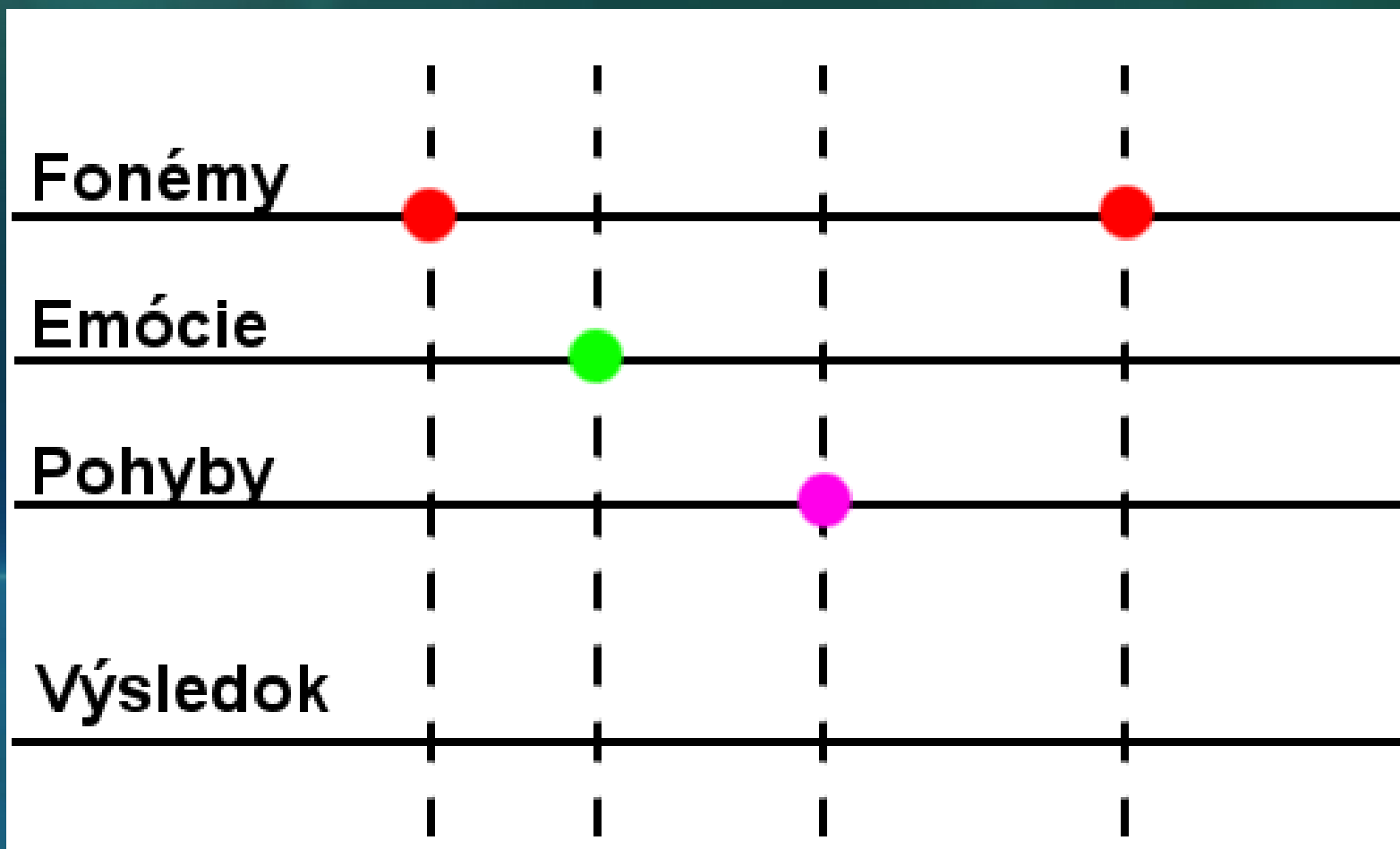


Príprava kľúčových snímkov

- Spracovanie udalostí
 - vytvorenie kľúčových snímkov na 3 časových priebehoch podľa typu udalosti:
 - fonémy, emócie, pohyby
- Vytvorenie výsledného časového priebehu
 - obsahuje zjednotenie kľúčových snímkov z časových osí všetkých priebehov
 - ak daný priebeh neobsahuje v danom čase kľúčový snímok, doplní sa interpoláciou
 - do výsledného priebehu sa parametre jednotlivých priebehov skladajú

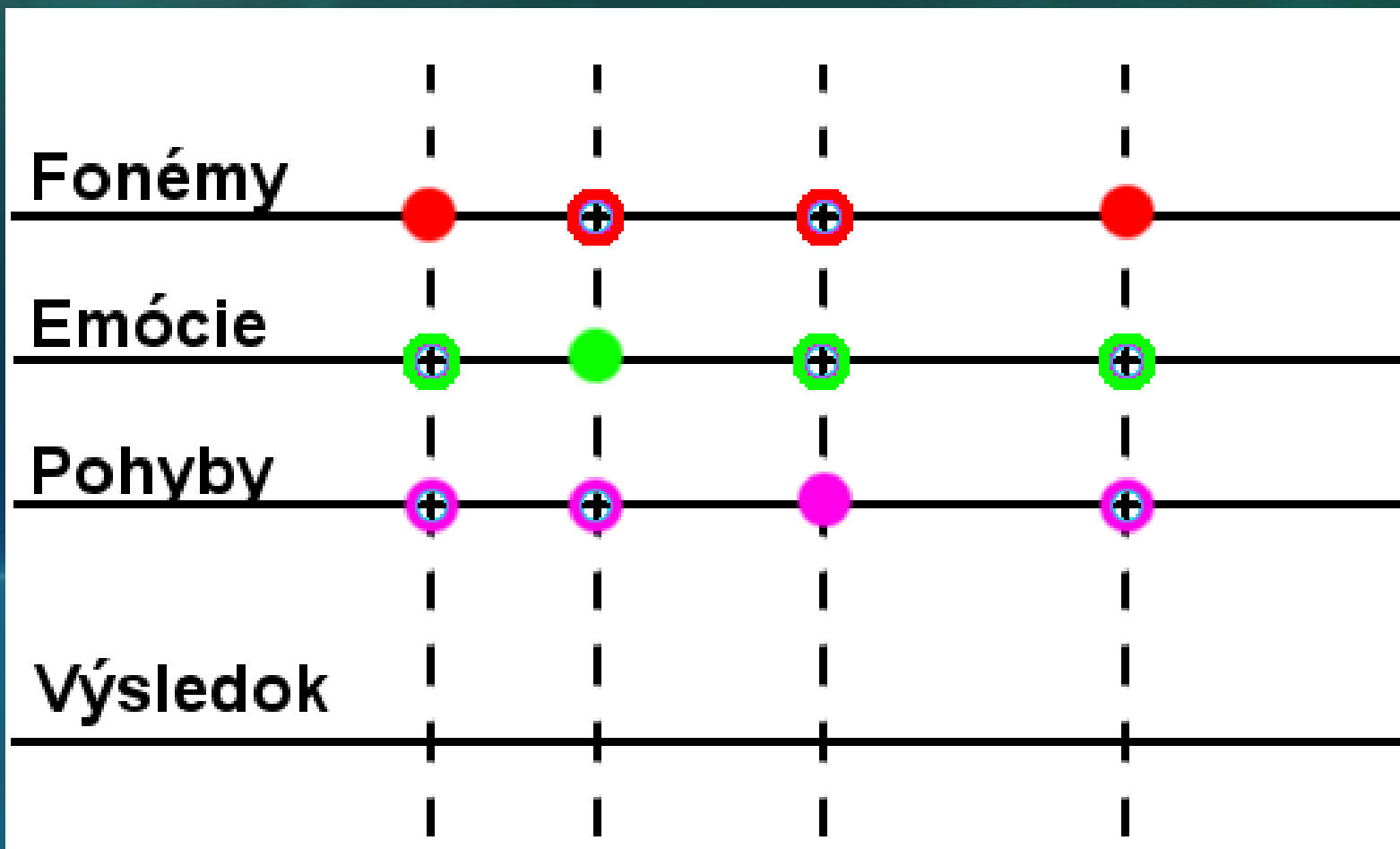
Príprava kľúčových snímkov

- Vloženie kľúčových snímkov z udalostí



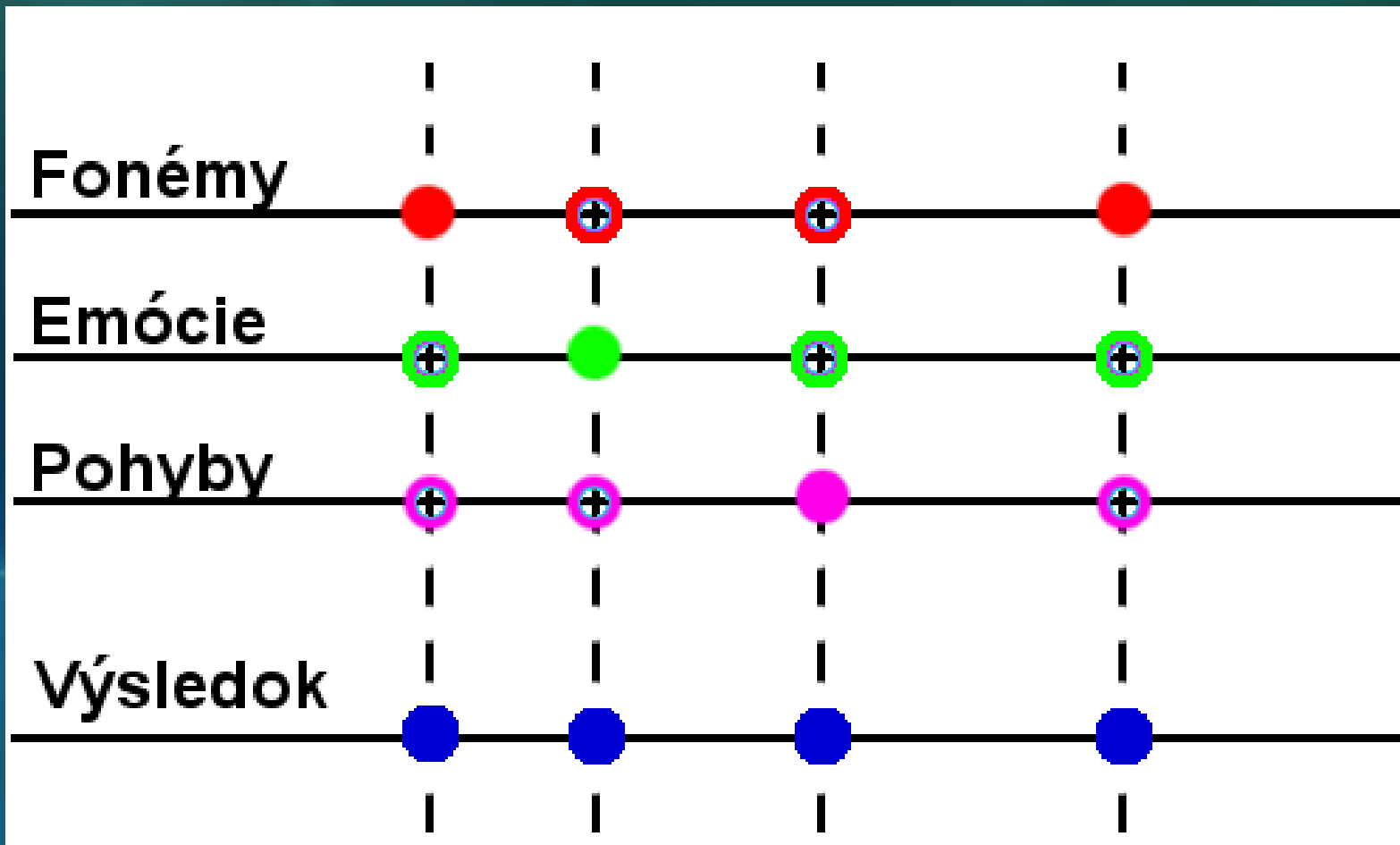
Príprava kľúčových snímkov

- Interpolácia chýbajúcich vizém

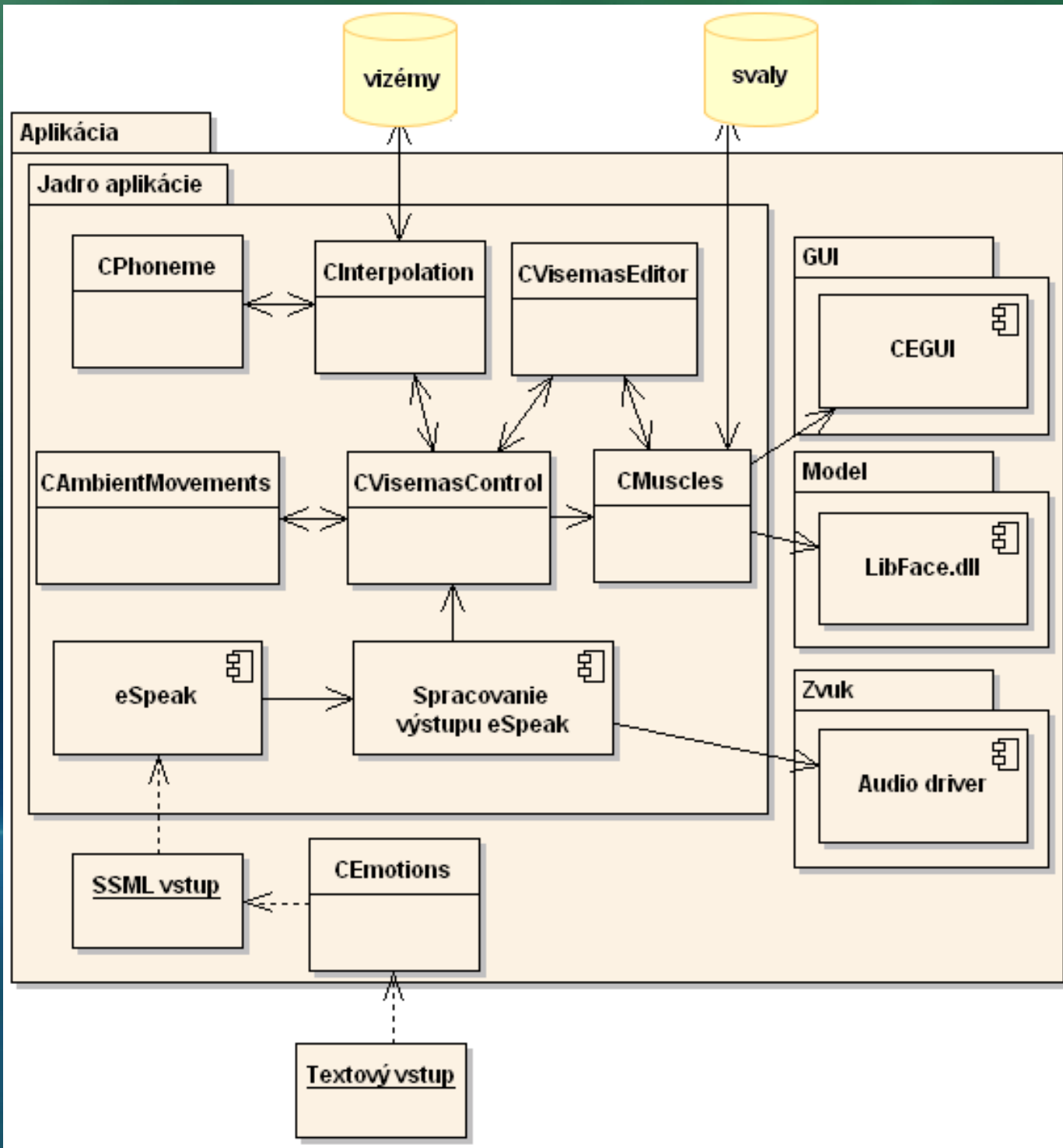


Príprava kľúčových snímkov

- Zloženie výsledného časového sledu vizém

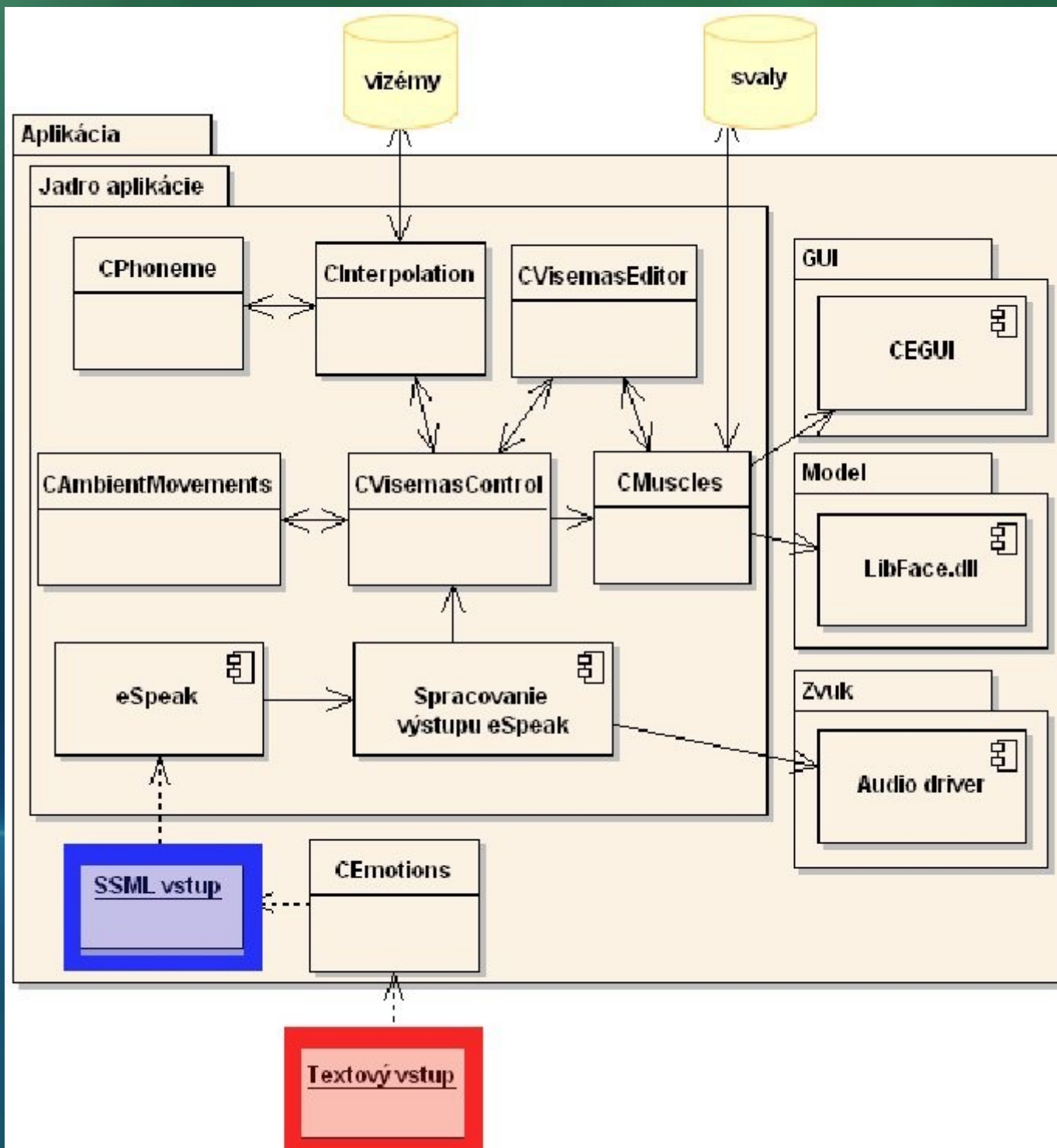


Tok údajov



Tok údajov

- Spracovanie do SSML
 - odstránenie smajlíkov
 - označenie interpunkcie



Tok údajov

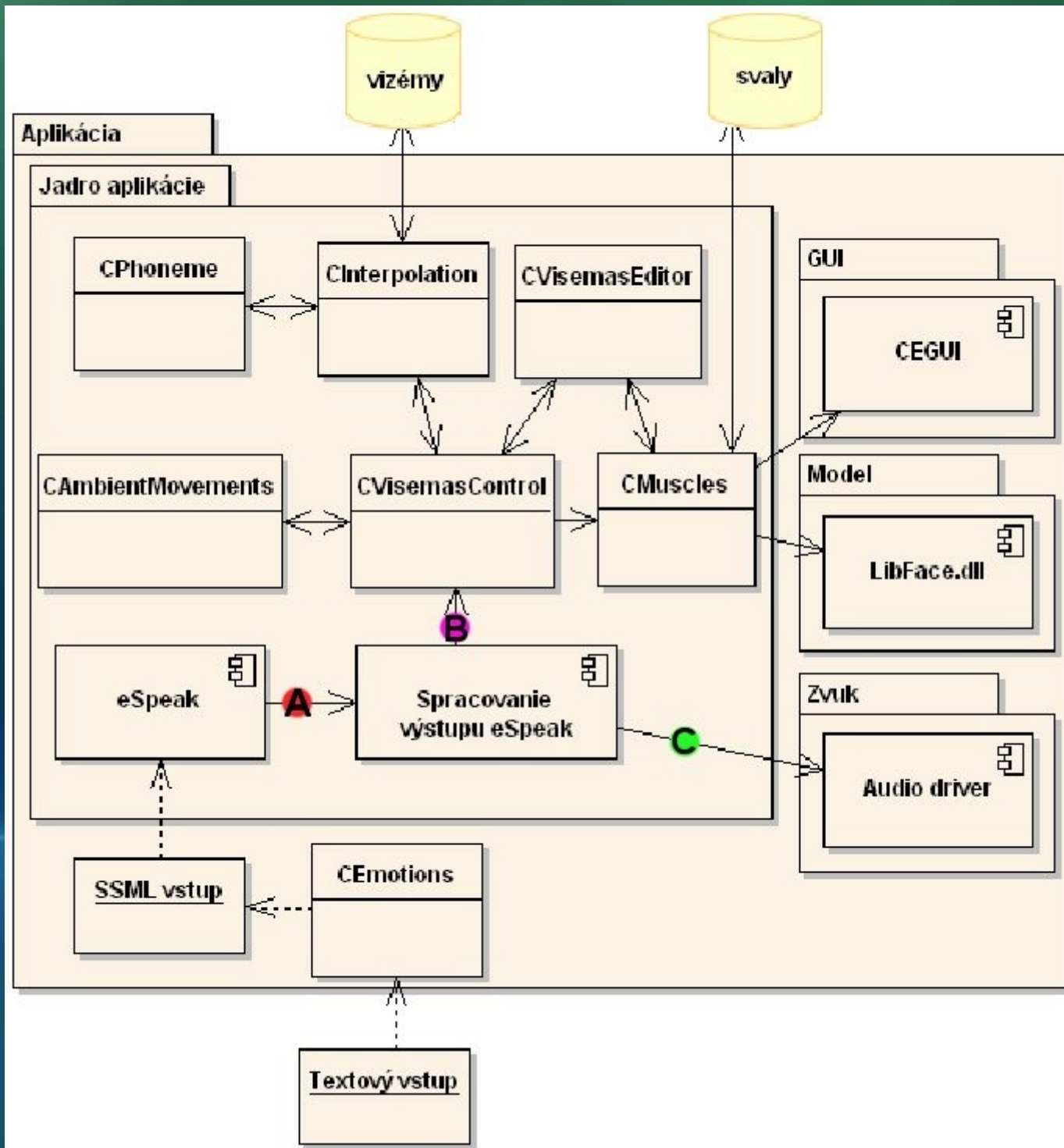
- Spracovanie výstupu eSpeak

Údaje:

A – udalosti,
audio

B – udalosti

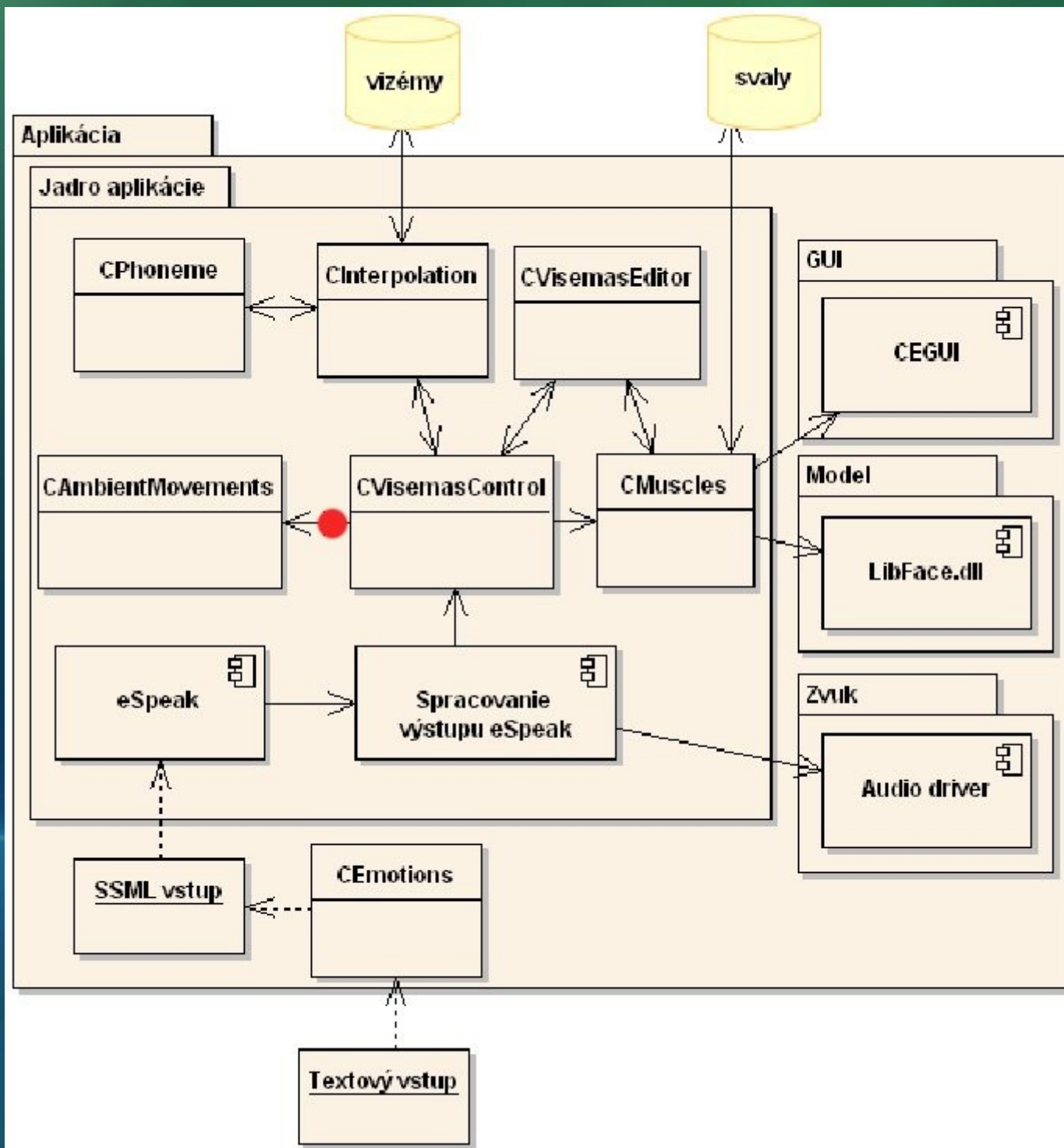
C – audio



Tok údajov

- Získanie pohybov pri reči

Údaje:
- udalosti vety



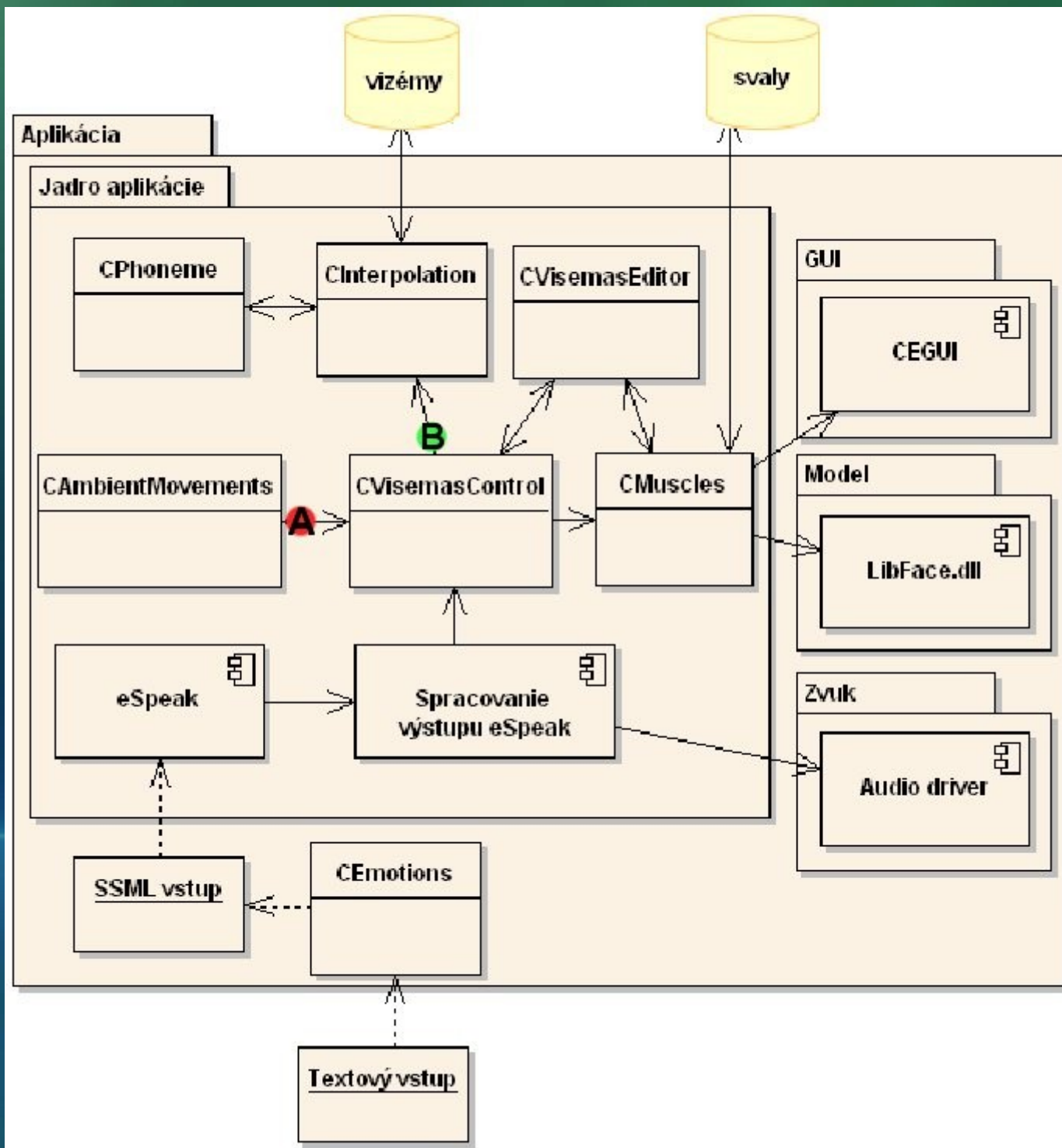
Tok údajov

- Získanie pohybov pri reči

Údaje:

A – mená pohybov

B – mená pohybov, emócií, foném

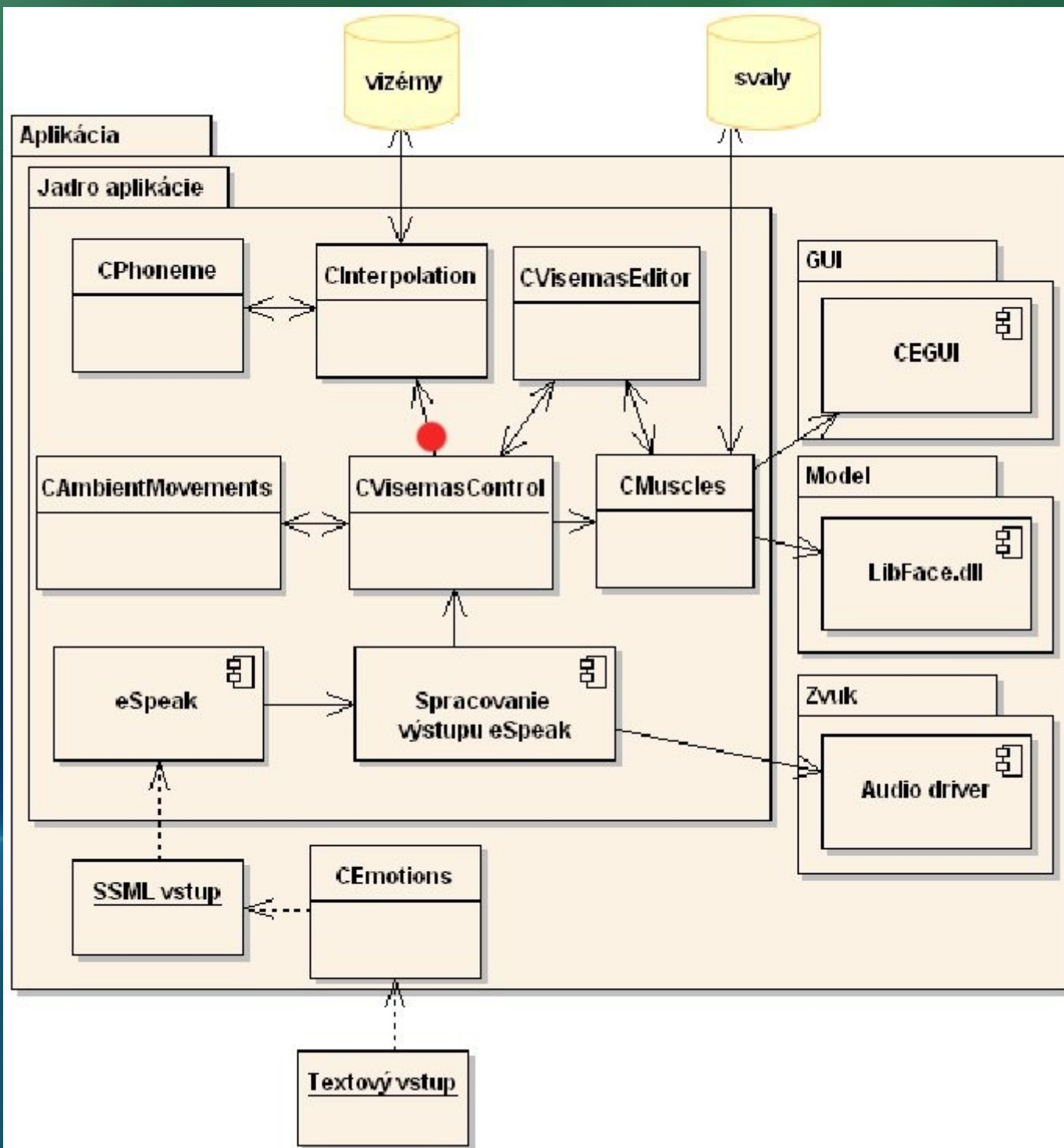


Tok údajov

- Získanie pohybov pri reči

Údaje:

- požiadavka s časovým údajom



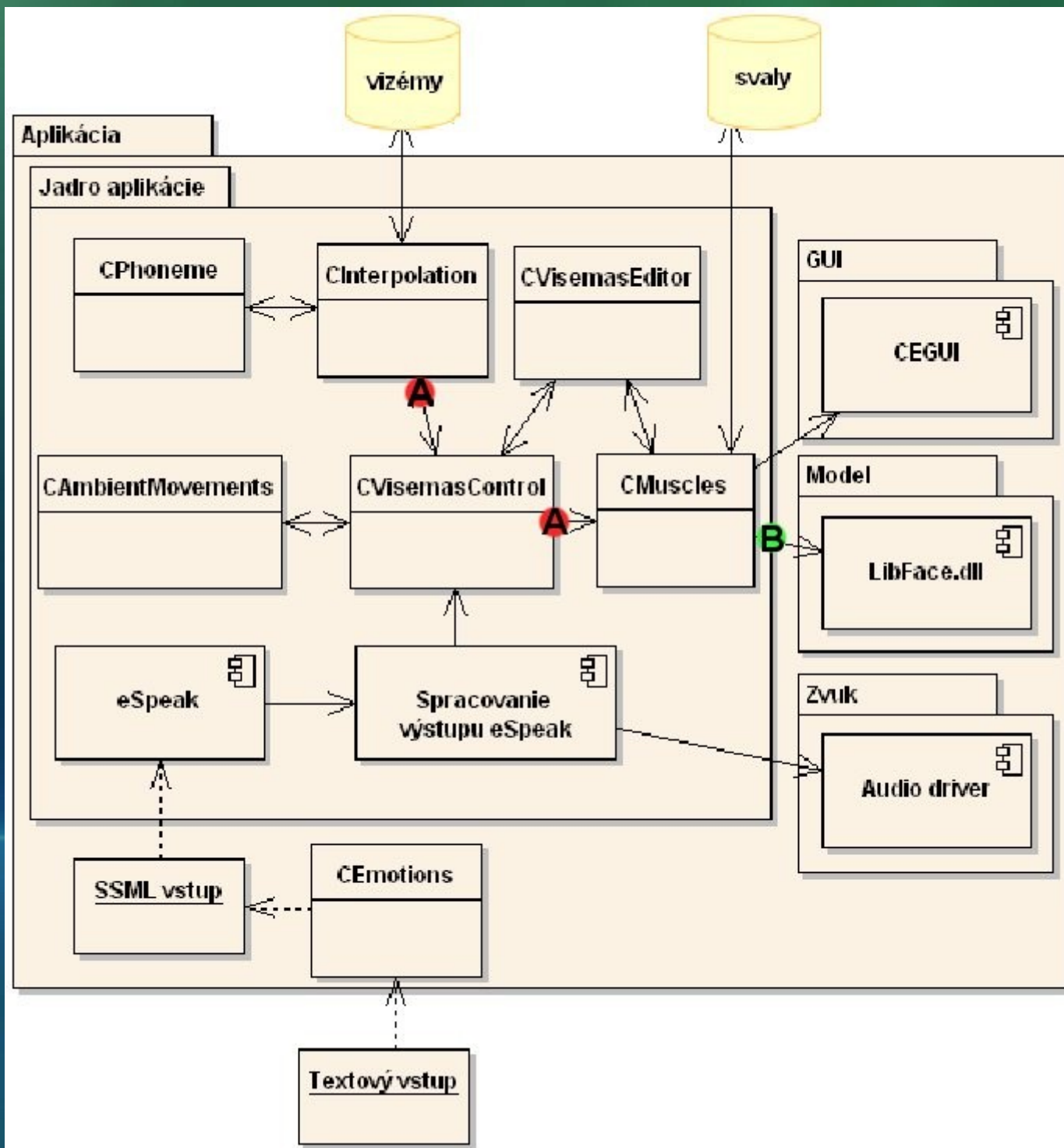
Tok údajov

- Získanie pohybov pri reči

Údaje:

A – výsledná vizéma

B – parametre svalov



Implementácia

- Implementácia
 - jazyk C++
 - prostredie MS Visual Studio .NET
 - využitie ATL knižníc
- Použitie existujúcich komponentov
 - model hlavy – knižnica LibFace
 - analyzátor SSML a syntetizátor reči – eSpeak
 - grafické používateľské rozhranie – CEGUI

Implementácia a overenie riešenia

- Synchronizácia reči
 - systémový čas
 - synchronizácia vláken
- Overenie riešenia
 - pozorovaním
 - pomocou záznamov v súbore log.txt
 - časovanie
 - aplikácia vizém

Výsledky práce

- Aplikácia

- synchrónna animácia modelu ľudskej hlavy
- demonštrácia riadenia modelu pomocou parametrov kontrakcie svalov modelu
- riadená pomocou XML jazyka (SSML)
- prezentácia praktického použitia knižnice LibFace



Ďalšie možné vylepšenia

- rozšírenie
 - o interpolácie a spracovanie foném
- možnosť modifikácie
 - vizém
 - rozsahu svalov
 - počtu a umiestnenia svalov
- zavedenie databázy
 - databáza modelov, rozloženia svalov, vizém
 - zdieľanie a paralelná práca na databáze

Ďakujem za pozornosť