

Werkplan Masterproef

Titel: Real-time signaal synchronisatie met accoustic fingerprinting

Bedrijf/onderzoeksgroep

Naam: IPEM

Tel: +32 (0)9 264 4107

Promotor(s):

E-mailadres: joren.six@ugent.be

Bestaande situatie en probleemstelling

Onderzoekers van het IPEM combineren vaak data van allerhande sensoren met video en audio-opnames om de relatie tussen muziek en beweging te onderzoeken. Om synchronisatie mogelijk te maken tussen die verschillende datastromen wordt momenteel de volgende methode gebruikt: voor elke datastroom wordt er synchroon een audio-opname gemaakt van het omgevingsgeluid. Aangezien dit omgevingsgeluid voor elke datastroom gelijkaardig is wordt het probleem van onderlinge synchronisatie tussen datastromen vereenvoudigd tot synchronisatie van de bijhorende audiostromen. Door acoustic fingerprinting technieken toe te passen kunnen audiostromen efficiënt en betrouwbaar met elkaar gesynchroniseerd worden. Zelfs al verschillen de audiostromen sterk in kwaliteit. Er is een implementatie van deze methode beschikbaar, maar de huidige implementatie laat echter enkel synchronisatie als nabewerking toe. Voor de eindgebruiker (die vaak niet technisch onderlegt is), is het lastig en tijdrovend om de synchronisatie telkens als nabewerking te moeten uitvoeren. Daarom is het de bedoeling om deze nabewerking overbodig. Wat precies de bedoeling is kan u lezen in de doelstellingen van het project.

(Deze tekst komt grotendeels toch van het voorstel op Plato, maar ik kan het probleem echt niet beter verwoorden dan hoe het daar omschreven staat.)

Doelstelling van het project

Het mogelijk maken om in real time gegevensstromen te synchroniseren met behulp van audio fingerprinting. Deze realtime implementatie zal geïmplementeerd worden als een Max/MSP module zodat de toepassing eenvoudig te gebruiken is. Het is belangrijk dat mensen met beperkte IT kennis deze toepassing ook kunnen gebruiken. Het is de bedoeling dat de verschillende datastromen, via het patch systeem, als invoer naar de zelf geschreven Max/MSP module worden gestuurd. Deze module zal er voor zorgen dat de verschillende datastromen worden gesynchroniseerd en dat het resultaat verder kan worden gebruikt in Max/MSP. Een mogelijkheid is dat het resultaat in WAVE formaat wordt teruggegeven met per kanaal een datastroom (in het eerste kanaal het audio signaal). Een andere module van Max/MSP kan dan zorgen voor het wegschrijven naar een WAVE-bestand.

Werkplan

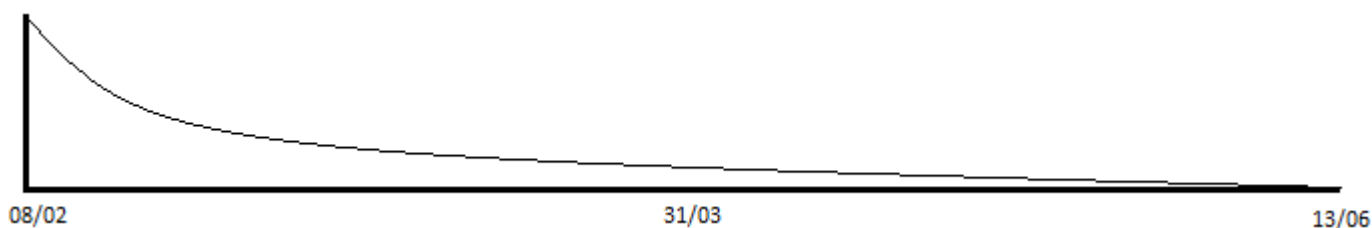
Aangezien de uitwerking van het probleem zich moeilijk laat onderverdelen in kleine deeltaken die week per week kunnen worden ingepland, heb ik, in overleg met mijn promotor, er voor gekozen om dit anders weer te geven als in het sjabloon van het werkplan.

In onderstaande diagrammen is de masterproef onderverdeeld in 4 grote delen: literatuurstudie, ontwerpen en programmeren van de software en algoritmen, schrijven van de thesis en het inwerken in de Max software.

Per diagram wordt er een ruwe schatting weergegeven van de hoeveelheid tijd ik hieraan zal spenderen. Dit in functie van waar we zitten in het tweede semester.

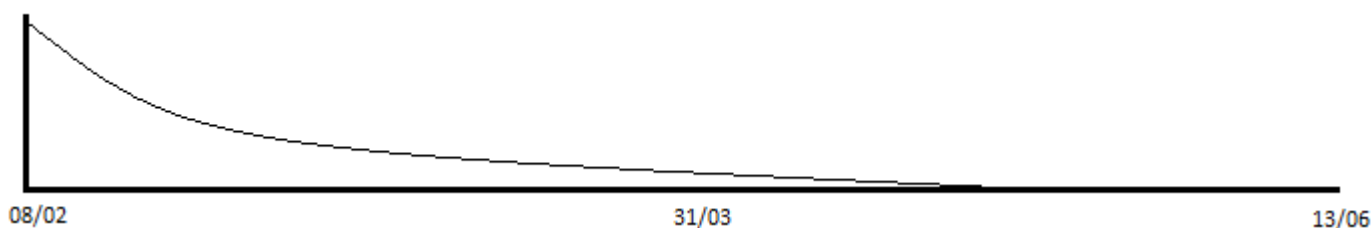
Vervolgens zal ik per deel de belangrijkste onderdelen ervan opsommen en toelichten.

Literatuurstudie



Het lezen en bestuderen van vakliteratuur in verband met audio fingerprinting en alles wat ermee te maken zal het volledige semester noodzakelijk zijn. Ook tegen het einde van de masterproef kan ik nog steeds zaken tegenkomen die mij niet duidelijk zijn en waarover ik extra opzoekingswerk moet doen. Het spreekt uiteraard voor zich dat ik mij hier toch vooral in het begin op zal moeten focussen. Onderaan dit document bevindt zich een lijst met literatuur die ik van plan ben door te nemen.

Inwerken in de software



Om de toepassing te kunnen ontwikkelen zal ik gebruik moeten maken van enkele softwarepakketten. Tot nu toe heb ik hier nog geen ervaring in, dus het zal zeker tijd kosten om hier mee te leren werken. Ook dit zal vooral tijd in beslag nemen tijdens het begin van de masterproef. De softwarepakketten die ik nog niet voldoende beheers zijn de volgende:

- Max/MSP: dit is de software waarin ik de module zal schrijven die er voor zal zorgen dat de verschillende streams in realtime gesynchroniseerd zullen worden.
- Panako: dit is de bibliotheek die er voor zorgt dat audiosignalen gesynchroniseerd kunnen worden, deze bibliotheek zal worden aangeroepen vanuit de Max/MSP module.

Onderaan dit document staan er enkele links met referenties die ik kan gebruiken bij het inwerken in de software.

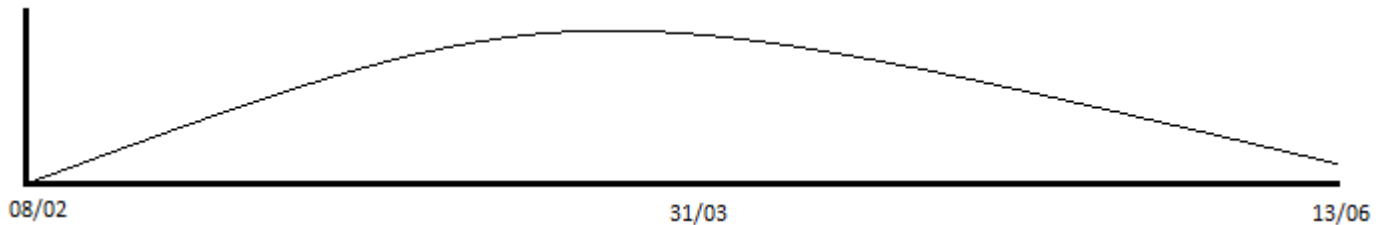
Het inwerken in de Max software zal ik als volgt aanpakken: Ik ga proberen om een zeer eenvoudige applicatie te schrijven gebruik makend van dezelfde architectuur die ik ook zou gebruiken voor de ontwikkeling van de module voor het real time synchroniseren. Deze architectuur zal bestaan uit de volgende componenten:

- Een onafhankelijke module die zorgt voor de verwerking van de gegevens en niet afhankelijk is van Max datatypes, libraries,...
- Een interface die methodes voorziet voor de aanvoer van de gegevens aan de onafhankelijke module.

- Twee implementaties van deze interface: Ten eerste: een implementatie die de verbinding met Max mogelijk maakt en er dus voor zorgt dat de toepassing bruikbaar is als Max module. Ten tweede: een test implementatie die niets met Max te maken heeft maar die ook gegevens kan aanleveren.

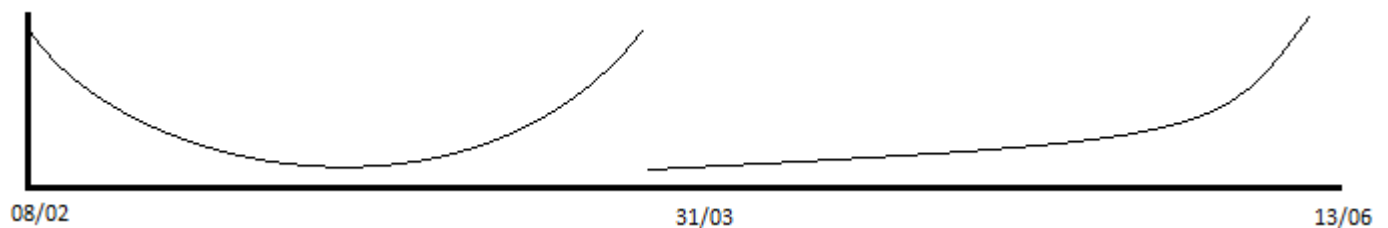
In de eerste zeer eenvoudige applicatie zal de onafhankelijke module bijvoorbeeld enkel iets uitschrijven. In het uiteindelijke resultaat is het de bedoeling dat deze module alle datastreams binnenkrijgt, deze synchroniseert met behulp van Panako, en vervolgens teruggeeft als resultaat.

Ontwikkelen



In het midden van het tweede semester zal de focus vooral liggen op het schrijven van de toepassing zelf. Dit is het moment dat ik al redelijk goed ben ingewerkt in Max/MSP en Panako, en dat ik al vrij veel kennis heb opgedaan door het lezen van de literatuur.

Schrijven van de thesis



Op drie verschillende momenten van het tweede semester zal de nadruk sterk liggen op het schrijven van de thesis. In het begin van het semester omdat ik op dat moment ook bezig ben met het doornemen van de vakliteratuur, met behulp van de kennis die ik hierbij opdoe zal ik in staat zijn om te beginnen schrijven aan de introductie.

Aangezien er op 31 maart een deadline is waarbij er 25 pagina's moeten worden ingediend, is het vanzelfsprekend ik mij ongeveer een week voor deze deadline hierop zal concentreren.

Ook tegen het einde van het tweede semester zal ik mij vooral focussen op het schrijven van de thesis. Uiteraard zal ik ook gedurende het hele semester hier aandacht aan besteden.

Door te nemen literatuur

- Paper over het synchroniseren van datastreams met behulp van audio fingerprinting: <http://0110.be/files/attachments/434/2015.synchronized-recording.pdf>
- Paper over hoe er kan worden omgegaan met tijdsverschillen en toonhoogte verschillen bij audio fingerprinting: http://0110.be/files/attachments/415/ismir_2014_panako_fingerprinter.pdf
- De originele paper van het gebruikte algoritme bij audiofingerprinting (het Shazam algoritme): <https://www.ee.columbia.edu/~dpwe/papers/Wang03-shazam.pdf>
- Introductiefilmpje over digitale audio: <https://xiph.org/video/vid1.shtml>
- Informatie over de opbouw van WAVE bestanden: <http://soundfile.sapp.org/doc/WaveFormat/>

Hulpmiddelen bij het inwerken in de software:

- Videotutorial reeks over Max/MSP: <https://www.youtube.com/playlist?list=PLVla8UkRzErsdnC1DEpDy7S9Ss7BhY7nY>
- Informatie over een Max/MSP module: http://0110.be/posts/TarsosDSP_PureData_or_MAX_MSP_external
- Broncode van PANAKO: <https://github.ugent.be/IPEM>