Multimedia Lab: Master’s Thesis Topics 2014-2015

Ghent University – iMinds
Faculty of Engineering and Architecture
Department of Electronics and Information Systems
Multimedia Lab
People

• Staff
  - Rik Van de Walle – senior full professor, head of MMLab
  - Peter Lambert – associate professor
  - Piet Verhoeve – 10% guest lecturer (iMinds)
  - Erik Mannens – project management & iMinds research lead
  - Jan De Cock and Wesley De Neve – project & research management
  - Ellen Lammens and Laura Smekens – administrative management
• 35 researchers
  - 50% PhD students
• Miscellaneous
  - ca. 15 thesis students per year
  - a few Summer internships each year
Research Activities (1/2)

• Cluster 1: **Video Coding** (Jan De Cock)
  - compression and transport of video
  - transcoding and scalable coding
  - high-dynamic range video

• Cluster 2: **Game Tech & Graphics** (Peter Lambert)
  - augmented and virtual reality
  - texture and mesh compression
  - path planning
Research Activities (2/2)

• Cluster 3: **Semantic Web** (Erik Mannens)
  - multimedia and interactivity on the Web
  - knowledge representation and reasoning
  - (big) data analytics and visualization

• Cluster 4: **Social & Visual Intelligence** (Wesley De Neve)
  - social media analysis
  - visual content analysis
  - deep machine learning
Standardization Activities

• **W3C** (World Wide Web Consortium)
  - new Web techniques
  - e.g., HTML5 and Media Annotations

• **MPEG** (Moving Picture Experts Group)
  - new compression techniques
    • e.g., H.264/AVC and 3-D Video Coding
  - new storage and transport techniques
    • e.g., MP4 file format and MPEG DASH

• **VQEG** (Video Quality Experts Group)
  - measurement of video quality
  - e.g., subjective quality evaluations
VIDEO CODING
Video Coding

Video coding aims at improving the visual quality of the video, while reducing the cost for both processing and delivery.

• Beyond-HD and more
  - High Dynamic Range
  - Multi-camera drones, 360 degree, DNA

• Processing
  - Encoding
  - Transcoding

• Optimizations
  - Machine learning for video processing
  - Quality based optimizations
High Dynamic Range (HDR):
- Optimize HEVC for HDR
- Transcode HDR to low dynamic range
Future inputs:

- Multi-camera drones
- 360 degree panorama video
- DNA
HEVC targets Ultra HD resolutions (4k, 8k):

- Complexity increases due to resolution and algorithmic design
- Hardware/software co-design
- Power aware encoding
Transcoding:
- Translate syntax
- Allow scalability
- Inter-standard transcoding

Syntax translation

Syntax translation + data re-use
Transcoding:
- Create compositions efficiently
- Extract region of interest
Machine learning for video compression

- Classify data to improve algorithms
- Off-line and online learning for encoding and transcoding
- Next generation still image coding
HTTP Adaptive Streaming:

- Network and device characteristics
- Optimize encoding levels and algorithms for multi-stream encoding
- Quality considerations
Contact Information

Zuiderpoort – office 3.04

Robin Bailleul
robin.bailleul@ugent.be

Johan De Praeter
johan.depraeter@ugent.be

Tom Paridaens
tom.paridaens@ugent.be

Luong Pham Van
luong.phamvan@ugent.be

Ruben Verhack
ruben.verhack@ugent.be

Niels Van Kets
niels.vankets@ugent.be

Sebastiaan Van Leuven
sebastiaan.vanleuven@ugent.be

Glenn Van Wallendael
glenn.vanwallendael@ugent.be

Jan De Cock
jan.decock@ugent.be
GAME TECH & GRAPHICS
Next-Gen Augmented Reality on Smart Eyewear

Impact of a limited FOV?:
- Compare models, APIs & SDKs
- Implement & test prototype(s)
- Improvements, user experience?

Full-view overlay rendering of AR?:
- Battery life?
- Processor drain?
- Workarounds?
  (sleep, cache, compr.)
- Feasible applications?

Recognition:
- Barcodes, QR codes, AR codes?
- Compare different APIs & SDKs
- Possible improvements
Streaming & Video Conferencing on Google Glass

HD video streaming on Google Glass

Google Hangout =
- Low resolution
- Low quality
- Frequent video jitter

What other solutions exist to enable:
- Live stream (point-to-point, one way)
- Video conferencing (two way)

How to offer the best HD quality experience?
- Streaming solutions (HTTP streaming)
- Using codecs currently not native
- Using more specific network protocols
Enhancing User Experience by Data Processing using Wearable Computing Devices

Wearable devices, new challenges:

- IN: New types of personal activity data
- OUT: What to communicate?
- OUT: How to communicate?
  - Limited display real-estate
  - Data processing is expensive!
    - Battery life
    - Processing power is limited

Identify types of data
Identify types of data processing
  i.e. what to offload to server

Efficient communication (sleep, bursts, proximity)
Privacy concerns
Grid-based Real-Time Multi-Agent Path Planning for Crowds in Dynamic Environments

Main novelty:
• Reuse internal path information (e.g., cached, hierarchical)
• Control over search horizon per agent

Requirements:
• Grid-based, arbitrary sized agents
• (sub-)linear scaling (1Mo+ agents)
• Collision avoidance
• Benchmark analysis

Based on Local vs. Global solver (e.g., potential fields, crowd simulation,.. )
• CPU (1 student)
• CPU & GPU (2 students)
Adaptive Collision Meshes for Dynamic Point Cloud Data Sets in Unity

Problems:
- It takes time to construct collision meshes
- Data is constantly updated

What we want:
- Make collision meshes adaptive as new cloud point data is discovered

Take into account:
- Types of agents: walking, flying, driving, ...
- Different agent collision sizes
- Changes based on events (traffic accidents), changes based on topology (elevators, bridges, ...)

Unity 3D implementation
Automatic Layout and Generation of 3D Worlds

Problem:
• Large 3D worlds are costly to build
• Vibrant 3D worlds require complex ontology and rule-sets (mistakes)

Goal:
• Procedurally generated 3D cities (outdoor / indoor) that can be densely populated

Requirements:
• Ontology rich (e.g., smart tags, reasoning)
• Artist assisting & augmenting
• Locally editable
• Taking into account physics laws, design principles and domain specific knowledge (i.e., architectural)
Optimizing Mesh Simplification through Voxelization

**Problem**
- Very high quality assets
- Simplification:
  - Pay artists to do the same thing again
  - Or, find a more automatic solution

**Result**
- Fast way to reduce the polygon count
- Making use of GPU power

**Skills**
- Mesh representations
- GPGPU programming
Texture Generation for Simplified 3D Models

Problem
• Simplified mesh shows aliasing
• We do not want to increase polygon count

Result
• Textures give the illusion of complexity
• Fewer polygons, without visual difference

Skills
• Offline rendering vs. real-time rendering
• GPU programming
Scalable Coding of Meshes

Bandwidth?

GPU processing power?
Scalable Coding of Meshes

Problem
• Storage, transmission and rendering of high-resolution models
• Differences in capabilities requires multiple versions of a model

Result
• Scalable representation and coding of a mesh
• Compact, lossless coding

Skills
• Signal processing
• Mesh representation
Contact Information

Jonas El Sayeh Khalil
jonas.elsayehkhalil@ugent.be

Ignace Saenen
ignace.saenen@ugent.be

Peter Lambert
peter.lambert@ugent.be

Zuiderpoort – office 2.29
SEMANTIC WEB

Enabling new applications through machine-understandable information.
Semantic Web topics

Linked Data Fragments
Digital Publishing
Visualizing the Web of Data
Storytelling with Linked Data
...
The current way of querying Linked Data does not work on a Web scale.
Servers only solve simple questions, clients solve complex questions.
Linked Data Fragments

query the Web — theses

query existing HTML pages
multi-source querying
querying videos and images
trusting query results
querying transport data

...
Digital Publishing

➢ From digital to print: closing the gap between digital and analog publishing

➢ Automatic Composition of Context-based Content in Digital Books

➢ The Machine-Understandable Book

➢ Normalization and Enrichment of HTML for machine-understandable websites
Digital Journalism

➢ Automatic **Newsworthiness** Assessment of Journalistic Content
Visualizing the Web of Data

How to visualize graph structured data?
More importantly: how do we interact with it?
Storytelling with Linked Data

Discover relations and paths in Linked Data

Identify, compare, associate optimally for target users

generate coherent stories in sets of (heterogeneous) data sources

apply to domains such as
- social news
- non-trivia fact finding
- research & academia
Personalised and Dynamically Built Monitoring Dashboards

**Goal:**
Investigate how to dynamically build dashboards visualising mash-up data interesting for the user

**Possible use cases:**
- Smart home surveillance
- Condition monitoring for offshore wind turbines
Semantic Web — come chat with us

Ruben
ruben.verborgh@ugent.be

Miel
miel.vandersande@ugent.be

Tom
tom.denies@ugent.be

Pieter
pieter.colpaert@ugent.be

Laurens
laurens.devocht@ugent.be
SOCIAL & VISUAL INTELLIGENCE
Natural Language Processing for Microposts using Deep Learning

Microposts:

Challenges:
- Deal with the Twitter slang
- Find the topic of the tweet (What is it about?)
- Detect the entities (What are the keywords?)

Techniques:
- Deep learning
- External Data Sources
- Social Network Information
Deep Learning for Entity Specific Sentiment Analysis of Reader Comments on News Reports

Abandon ship? In recent maritime disasters, captains don't hang around

Comments disagreeing with journalist

Challenges: How to measure this difference? Which techniques are suitable?
Automatic Product Recognition in YouTube Videos and Subsequent Linkage to eBay
Categorization of Short-form Mobile Video

Basketball
Stephen Curry
NBA
Sports
Basketball
Stephen Curry

#crossover2013 #crossover
Stephen Curry #comedy #loop
#KevinHart #nba #basketball
Aug 26 2013
Video Object Tracking

Recognize
Track
Label

zentrick
YouTube
Linking Twitter Messages to TV Video Fragments

Link tweets to TV fragments

😊 or 😞?
An Automatic, Personalized Sports Commentator by Making Use of Deep Learning and Social Media
Semantic Hashing for Near-Duplicate Video Clip Detection on Online Platforms for Video Sharing

• Problem statement
  – YouTube and Instagram contain a lot of video content
    o visual redundancy
    o infringement intellectual property

• Goal
  – detection of video copies by means of deep learning techniques
    o translation of video content into so-called semantic hashes
    o search for similar semantic hashes
Of LFC pressure, Arsenal have the ball. Arsenal can't keep it though, LFC are closing them down.

Arsenal Football Club is an English Premier League football club based in Holloway, London. One of the most successful clubs in English football, Arsenal have signed for a big club. Are you a big club? Where's your European Cups?

Liverpool Football Club is an English Premier League football club based in Liverpool. The club has won eighteen League titles, seven FA Cups, and numerous other domestic trophies. It's Liverpool, and they're a big club. This is why we want them to lose. arsliiv.

We can win this.
Geo-Annotation of Road / Terrain Type by Participatory Bike-Sensing (+ Web-based Geo Enrichment)

Web-based GEO-enrichment gives more flexibility and freedom to the user.
Automatic Geo-Mashup Generation of Outdoor Activities (+ POI 2.0)

GEO-MASHUP = online geotagged media resources related to the “geographic keypoints” where the outdoor activity was performed.

POI 2.0 → time-dynamic, personalized POI

http://multimedialab.elis.ugent.be/geomashup/
Touristic Digital Dissemination of the Flemish Waterways

1) (Semi-)automatic collection and management of **geo-tagged media** on waterways

2) Generation of contextualised **recommendations**

3) Archiving of geo-tagged sensor data on the boat/water conditions (**on-/off-boat data**)

Greetings from Ghent, Belgium

PANORAMIO results

DBPEDIA entities

http://en.wikipedia.org/wiki/Saint_Nicholas%27s_Church%2C_Ghent
http://dbpedia.org/resource/Saint_Nicholas%27s_Church%2C_Ghent
http://dbpedia.org/resource/Saint_Bavo_Cathedral
http://dbpedia.org/resource/Saint_Nicholas%27s_Church%2C_Ghent
http://dbpedia.org/resource/Oostakker_Basilica

FLICKR results
Fire Detection using Visual and Social Sensors

Sensor-based monitoring of fire development
Detection and localisation of forest fires by means of social media

1) Automatic detection and localization of flames and smoke based on complementary sensors (preferably video + other volume sensors)
2) Optimization of the fusion of sensor data
3) (Visual) representation of the fire development, facilitating the interpretation of the sensor data + evaluation of a real-world test case

Multimodal Recognition of Human Actions
Contact Information

Abhineshwar Tomar
abhineshwar.tomar@ugent.be
@abhineshwar

Baptist Vandersmissen
baptist.vandersmissen@ugent.be
@BaptistV

Steven Verstockt
steven.verstockt@ugent.be

Fréderic Godin
frederic.godin@ugent.be
@frederic_godin

Viktor Slavkovikj
viktor.slavkovikj@ugent.be

Wesley De Neve
wesley.deneve@ugent.be
@wmdeneve

Zuiderpoort – office 2.24
ADDENDUM
Semantic Web Technologies for Natural Language Understanding

Dr. Gaëtan Martens
gaetan.martens@nuance.com
Nuance Automotive R&D

April 2014, MMLab -UGent
Natural Language Understanding (NLU)

In any given language (therefore, in NLU), there are virtually unlimited ways of speaking expressions to activate a function:

- I want to call Bob on his cell ................
- How do I call Bob on his iPhone?................
- Can you buzz Bob for me on cell?............... 
- Give Bob a call on his cell............................
- Select Bob and then cell and put me through.....
- Find Bob in my contacts and place a call for me on his mobile phone.................................
- I need to talk to Bob. Get him on his cellphone...
- I meant to call Bob on his cell, do I have his number and can you give him a ring?...............
Natural Language Understanding (NLU)

- I want to call Bob on his cell .........................
- How do I call Bob on his iPhone?........................
- Can you buzz Bob for me on cell?........................
- Give Bob a call on his cell................................
- Select Bob and then cell and put me through...........
- Find Bob in my contacts and place a call for me on his mobile phone................................................
- I need to talk to Bob. Get him on his cellphone......
- I meant to call Bob on his cell, do I have his number and can you give him a ring?.......................
Natural Language Understanding (NLU)

- I want to call Bob on his cell
- How do I call Bob on his iPhone?
- Can you buzz Bob for me on cell?
- Give Bob a call on his cell
- Select Bob and then cell and put me through
- Find Bob in my contacts and place a call for me on his mobile phone
- I need to talk to Bob. Get him on his cellphone
- I meant to call Bob on his cell, do I have his number and can you give him a ring?

\[ \text{Call}(\text{Bob};\text{cell}) \]
Problem description
- Natural Language Understanding (NLU) -

In order to recognize the functionality, one must model as much as possible what users say to activate the functionality

This modeling includes the use of trained classifiers which:
- Correctly recognize previously collected/modeled sentences
- Correctly attribute previously unmodeled sentences

Quality of models depend mainly on the data quality, but:
- Manual annotation takes time, is a cumbersome task, and is error-prone
- Data are gold: reuse (possibly previously annotated) data
- Data can be ambiguous
Semantic Web Technologies for NLU

Goal
- Facilitate the data annotation process
- Process results in a more consistent annotation
- Reuseable data (no re-annotations required)
- Obtain “cleaner & better” data for training and testing models

Rely on Semantic Web technologies
- For data annotation
- To express the semantics in the data
- Formally defined concepts, properties, and relationships
- Ontologies!

OWL
Semantic Web Technologies for NLU

Possible directions / tasks for this thesis
- Develop new ontologies, reuse, or adapt existing ones
- Develop/extend annotation tool which relies on ontologies (Software Development)
- Define and measure annotation quality (i.e., consistency)
- Annotate data (data obtained from Nuance)
- Train and test models using annotated data
- Effect of annotation quality on trained models