Analysing Daily Activity Logs for Smart Interaction

TRAN Minh-Triet
University of Science
Vietnam National University-Ho Chi Minh city
❖ **Introduction:**
  - Smart Environment and Smart User
  - Daily Activity Logs

❖ **Analyzing Daily Activity Logs for Smart Interaction:**
  - Augmented data/services by recognizing the current context and retrieving similar known cases.
  - Find lost items, retrieve or verify memories by searching daily logs.
  - Reminiscence can help people to positively revive past memories and connections with their relatives.
  - Detect regular events and anomalies from surveillance systems or sousveillance archives for appropriate actions.
  - Event simulation in virtual or mixed reality environments can be generated from real life data for education and training.

❖ **Conclusion**
April 15, 2018 at Hoan Kiem lake

April 15, 2018, at Bach Ma (White Horse) Temple

April 15, 2018 near Hoan Kiem lake

May 8, 2018 in Old Quarter, Hanoi
Introduction

Smart City

- Provide smart features, utilities, services
- Collect & process information
- Provide infrastructure & support for further development

Smart Citizen

- Exploit smart features, utilities, services
- Contribute information
- Develop utilities, ecosystem, community

Smart Environment

Smart User

- Look
- User
- Feel
- Usage
Daily Activity Logs

Official/Public Systems

Surveillance

Personal/Private Systems

Sousveillance

Source: Internet
Wide variation of **data types** and **data sources**

- Visual data
- Audio data
- Text data
- Information from various sensors
- Personal biometrics
- Human activities
- …
Multi-disciplinary Field of Research

- **Multimedia Analytics**: A variety of data, different timings, different accuracies, needing different tools.
- **Information Retrieval**: Scalable & efficient indexing with contextual querying and no defined unit of retrieval.
- **Human Computer Interaction**: Develop fixed and ubiquitous capture & access methods for all stakeholders.
- **Pervasive Computing**: Use-cases need pervasive access and contextual querying.
- **Ethics & Privacy**: The ethics of how to use rich personal data & doing so in a privacy-aware manner.
- **Ethnography**: The customs of individual peoples and cultures.
- **Memory**

Augmented data/services by recognizing the current context and retrieving similar known cases
Landmark Recognition for Tourism with Visual Instance Retrieval, HCII 2016
Traditional & Common Approach

Feature encoding

Bag of Visual Words

BoVW Framework for Visual Retrieval

BoVW Framework - Indexing

Feature extraction
- Hessian affine detector
- RootSift

Clustering
- Approximate K-mean

Quantization
- Natural soft-assignment
- Burst weighting

Build inverted index

Database-side feature augmentation

Visual-word vectors

tf-idf weighting

Database

James Philbin, Ondřej Chum, Michael Isard, Josef Sivic and Andrew Zisserman, Object retrieval with large vocabularies and fast spatial matching, CVPR 2007
Relja Arandjelović and Andrew Zisserman, Three things everyone should know to improve object retrieval, CVPR 2012.
BoVW Framework for Visual Retrieval

BoVW Framework - Retrieval

Feature extraction: Hessian affine detector, RootSIFT
Quantization: Natural soft-assignment, Burst weighting

Query Image

Query expansion: Average query expansion
Spatial verification: NoSAC

Query Image

Similarity: Cosine similarity
Rank list

Rank list

1\textsuperscript{st} 2\textsuperscript{nd} 3\textsuperscript{rd} 4\textsuperscript{th}

James Philbin, Ondřej Chum, Michael Isard, Josef Sivic and Andrew Zisserman, Object retrieval with large vocabularies and fast spatial matching, CVPR 2007
Relja Arandjelović Andrew Zisserman, Three things everyone should know to improve object retrieval, CVPR 2012.
Landmark retrieval and recognition: Challenging Problem!

https://www.kaggle.com/c/landmark-retrieval-challenge

https://www.kaggle.com/c/landmark-recognition-challenge
Who am I?
Smart Identity-based Authentication
Face Authentication

Conventional CNN approach?

User Authentication with Facial Features Using Convolutional Neural Networks for Smart Interactive Environments, Nhan Dam, Vinh-Tiep Nguyen, Minh N. Do, Anh-Duc Duong, Minh-Triet Tran, ISVC 2015
Domain Adaptation?

Illuminations

Aging

Camera parameters

Making up
Adversarial adaptation for face recognition

Goal: to learn $M_t$ so that the discriminator cannot distinguish the domain of a feature vector encoded by either $M_s$ or $M_t$.

What transformations? How to transform?
Face Recognition & Clustering for Image & Video Retrieval

- Character Flows in Video Sequence
- Person-based Video Navigation

Face Detection
Face Recognition
Person Re-identification

Multi-region Size CNN for Short-Period Gait-based Authentication

DeepSense-Inception for Multi-Period Gait-based Authentication

Khac-Tuan Nguyen, Thanh-Luong Vo-Tran, Dat-Thanh Dinh, and Minh-Triet Tran, Gait Recognition with Multi-Region Size Convolutional Neural Network for Authentication with Wearable Sensors, FDSE 2017
Ha V. Hoang, Minh-Triet Tran, DeepSense-Inception: Gait Identification from Inertial Sensors with Inception-like Architecture and Recurrent Network, CIS 2017
Personalized Smart Advertisement
Augmented Reality-based Shopping Assistance
Personalized Smart Shopping

Possible applications:

- Online community of customers
- m-commerce

Product/Book

Photo captured by camera

Features extracted by mobile device

Wiki articles
Multimedia objects

Landmark recognition service

Get wiki articles
Get video clips

Make purchase order

...
Food Recognition & Recommendation

Saliency Estimation

Main Area Extraction

Linear Classification

Pre-processing

High-level feature vector

Pre-trained CNN on large-scale dataset
Visual Attention for Object Detection

2 person(s)
Social Relation Trait Discovery from Visual LifeLog Data

Interpersonal Circle (Kiesler, 1983)

Tung D. Dinh, Dinh-Hieu Nguyen, Minh-Triet Tran, Social Relation Trait Discovery from Visual LifeLog Data with Facial Multi-Attribute Framework, INDEED 2017
Facial Multi-Attribute Framework: Siamese Structure

Facial Multi-Attribute Network → Facial Multi-Attribute Network

Facial Multi-Attribute Network → Facial Multi-Attribute Network

Facial Single-Attribute Network 1 → Facial Single-Attribute Network 2 → \[ a_1 \]

Facial Single-Attribute Network k → \[ a_k \]

Facial Single-Attribute Network

Social Relation Trait Prediction

Dominant
Competitive
Trusting
Assured

Tung D. Dinh, Dinh-Hieu Nguyen, Minh-Triet Tran, Social Relation Trait Discovery from Visual LifeLog Data with Facial Multi-Attribute Framework, INDEED 2017
Find lost items, retrieve or verify memories by searching daily logs
Where is my lost item?

When did I last see this item?
Where is my lost item?
Maybe I have seen this item, but when and where?
Lifelogging Data Pre-processing

- Lifelogging Data
  - Visual Clustering for Images
    - Visual Shot Detection
    - Similar Shot Linkage
  - Concept Extraction
    - Place Recognition
    - Action Detection
    - Entity Detection
  - Augmented Data Processing

Thanh-Dat Truong, Tung Dinh-Duy, Vinh-Tiep Nguyen, Minh-Triet Tran, Lifelogging Retrieval based on Semantic Concepts Fusion. ICLR 2018
Visual Clustering for Images

Frame $t$

Frame $t+1$

Magnitude of optical flow vectors (with FlowNet2.0)

Thanh-Dat Truong, Tung Dinh-Duy, Vinh-Tiep Nguyen, Minh-Triet Tran, Lifelogging Retrieval based on Semantic Concepts Fusion. ICLR 2018
Visual Shot Clustering

Input image

Lifelogging Visual Data

BoVW Visual Retrieval

Ranklist and distances

Visual Shot Distance

Visual Shot Connected Components

Thanh-Dat Truong, Tung Dinh-Duy, Vinh-Tiep Nguyen, Minh-Triet Tran, Lifelogging Retrieval based on Semantic Concepts Fusion. ICLR 2018
Visual Shot Clustering

August 15, 2016

August 16, 2016
Scene Category and Attributes

Location information
- Environment type
- Scene category
- Scene attributes

Single image
MIT Place API
Visual shot
Shot cluster

Thanh-Dat Truong, Tung Dinh-Duy, Vinh-Tiep Nguyen, Minh-Triet Tran, Lifelogging Retrieval based on Semantic Concepts Fusion. ICLR 2018
Scene Category and Attributes

(a) home_theater  (b) coffee_shop, fastfood_restaurant  (c) car_interior
“I am building a chair that is wooden in the late afternoon. I am at work, in an office environment, beside a yellow partition wall. There are plastic plants on the partition wall. Books and a trolley can be seen behind me on the ground. Since I am engaged in physical activities, my heart-rate has raised above 100bpm.”

Thanh-Dat Truong, Tung Dinh-Duy, Vinh-Tiep Nguyen, Minh-Triet Tran, Lifelogging Retrieval based on Semantic Concepts Fusion. ICLR 2018
“I am walking out to an **airplane** across the **airport apron**. I stayed in an **airport hotel** on the previous **night** before checking out and walking a short distance to the **airport**. The weather is very nice, but cold, with a clear blue **sky**. There is a **man** walking to the **airplane** in front of me with a **blue jacket**, **green shoes** and a **black bag**.”

Thanh-Dat Truong, Tung Dinh-Duy, Vinh-Tiep Nguyen, Minh-Triet Tran, *Lifelogging Retrieval based on Semantic Concepts Fusion*. ICLR 2018
Reminiscence can help people to positively revive past memories and connections with their relatives.
NowAndThen: Social Network-based Photo Recommendation for Reminiscence

Automatically retrieve visually similar photos for reminiscence
Automatically retrieve visually similar photos for personalized annotation and indexing.
Recommend personalized tags

New Photos -> Capture -> User’s Device -> Upload

Choose & Save

Share

Annotations

Social Network

User’s Device

Photos Database

Similar

Annotations Database

Our Server

Bao Truong, Thuyen V. Phan, Vinh-Tiep Nguyen, Minh-Triet Tran: Personalized Annotation for Photos with Visual Instance Search. HCI (3) 2016: 425-435
Learn to annotate photo from samples provided by users

Photo → Add Annotation → Photo & New Annotation → Upload

Choose & Save

Similar

User’s Device

Photos Database

Annotations Database

Our Server

Bao Truong, Thuyen V. Phan, Vinh-Tiep Nguyen, Minh-Triet Tran: Personalized Annotation for Photos with Visual Instance Search. HCI (3) 2016: 425-435
Detect regular events and anomalies from surveillance systems or sousveillance archives for appropriate actions.
Traffic Analysis

NVIDIA AI City Challenge, CVPR 2018

The video clips are from Track 1 of NVIDIA AI City Challenge, CVPR 2018
Challenges in Traffic Videos

Ideal scenario

Dark environment, bad illumination case

Shaking & blurring camera

Manually recorded video

The video clips are from Track 2 of NVIDIA AI City Challenge, CVPR 2018
Adaptive Vehicle Detector

Multiple Adaptive Object Detectors

- Annotate boxes
- Generate low-res data
- High-res training data

Train detector for front and back views
Train detector for side view
Train detector for tiny instances

Region-based Adaptive Set of Detectors

- Annotate boxes
- Generate low-res data
- High-res training data

(a) Input frame
(b) Optical flow for moving entities
(c) Blobs to estimate vehicle's size

Traffic Flow Analysis with Multiple Adaptive Vehicle Detectors and Velocity Estimation with Landmark-based Scanlines
Minh-Triet Tran, Tung Dinh-Duy, Thanh-Dat Truong, Vinh Ton-That, Thanh-Nhon Do, Quoc-An Luong, Thanh-An Nguyen, Vinh-Tiep Nguyen, and Minh N. Do, NVIDIA AI City Challenge, CVPR 2018
Velocity Estimation

Video with depth data?
Depth estimation for outdoor environment?

Moritz Kampelmuhler, Michael G. Muller, Christoph Feichtenhofer, Camera-based vehicle velocity estimation from monocular video, Vehicle Velocity Estimation Challenge (winner), CVPR2017
Context Analysis

Supermarket
How many counters should be opened?
- From history logs?
- From current situation?

Student Enrollment to Universities
How many counters should be opened?
- From history logs?
- From current situation?
Tourism/Attractions
How long should we wait at an attraction?
- From history logs?
- From current situation?

Source: Internet
Context Analysis

Power/ Air Conditioner Consumption
Estimate/Predict Power/ Air Conditioner Consumption
- From history logs?
- From current situation?

Source: Internet
Scene & event simulation in virtual or mixed reality environments generated from real life data
Do we perceive changes in moving velocities? A comparison between physically and non-physically powering the movement.

HCMUS 2018
Scene Synthesis with Semantic Segmentation

Instance Re-Identification Flow for Video Object Segmentation
Trung-Nghia Le, Khac-Tuan Nguyen, Manh-Hung Nguyen-Phan, That-Vinh Ton, Toan-Anh Nguyen, Xuan-Son Trinh, Quang-Hieu Dinh, Vinh-Tiep Nguyen, Anh-Duc Duong, Akihiro Sugimoto, Tam V. Nguyen, and Minh-Triet Tran,
Scene Synthesis with Semantic Segmentation

Instance Re-Identification Flow for Video Object Segmentation
Trung-Nghia Le, Khac-Tuan Nguyen, Manh-Hung Nguyen-Phan, That-Vinh Ton, Toan-Anh Nguyen, Xuan-Son Trinh, Quang-Hieu Dinh, Vinh-Tiep Nguyen, Anh-Duc Duong, Akihiro Sugimoto, Tam V. Nguyen, and Minh-Triet Tran,
RGB-D to CAD for 3D Virtual Scene Reconstruction/Generation

Real Partial 3D Fragments

Virtual Ideal 3D Models
RGB-D to CAD for 3D Virtual Scene Reconstruction/Generation

SHREC’18: RGB-D Object-to-CAD Retrieval

View sequence

![Images of rings with different colors and objects]

- Ring Classifier
- Saliency Analysis
- Prediction scores
- Confidence Evaluation
- Confidence score

Minh-Triet Tran, Ngoc-Minh Bui, Trong-Le Do, et.al. (Winner), EuroGraphics 2018
View sequence of the $i^{th}$ view-ring ($1 \leq i \leq 7$)

ResNet 50
ResNet 50
ResNet 50
ResNet 50
ResNet 50
ResNet 50
ResNet 50
ResNet 50

16384

LSTM

512 Fully Connected + ReLU
256 Softmax
256
20

RGB-D to CAD for 3D Virtual Scene Reconstruction/Generation

SHREC’18: RGB-D Object-to-CAD Retrieval
Minh-Triet Tran, Ngoc-Minh Bui, Trong-Le Do, et.al. (Winner), EuroGraphics 2018
Conclusion
Daily Activity Log Analysis

❖ **Answer:**
  - Find an item from the digital self
  - Validate a memory
  - Contextual support

❖ **Reflect:**
  - Quantified-Self Analysis
  - Self-discovery

❖ **Reminisce:**
  - Reminiscence Therapy
  - Social applications

❖ **Remind:**
  - Contextual Reminders

Multidisciplinary Problems and Approaches

Data Processing
A variety of data, different timings, different accuracies, needing different tools.

User Experience
Develop fixed and ubiquitous capture & access methods for all stakeholders.

Personal Data
The ethics of how to use rich personal data & doing so in a privacy-aware manner.

Multimedia Analytics

Information Retrieval
Search & Retrieval
Scalable & efficient indexing with contextual querying and no defined unit of retrieval.

Human Computer Interaction

Pervasive Computing
Anywhere, Anytime
Use-cases need pervasive access and contextual querying.

Ethics & Privacy

Ethnography

Memory

Collaboration in Various Fields of Research

Lifelog

- Citizen Engagement
  - Smart Citizen App
  - Smart Citizen Web Portal
  - Smart Citizen Chatbot
  - Smart Citizen Commute

- City Control
  - Control Centre
  - City BI
  - Citizen CRM

- Smart Citizen Integration
  - Smart Citizen integration Interface (SCII)
  - SAP Integration Consulting
  - MS CRM Integration Consulting
  - IOT Integration Consulting

- Field Agent
  - Field Agent App
Thank you for your attention