Activity Sensing Recognition

Wed - 1 to 2pm
TA: Rahi Kalantari

Summary + Critique
Required + Optional
Readings.

6 Homework →
4 Working with data.

Next Class
1. Panel of Experts (Inertial Sensing)
2. Reading Assignments
3. Your survey (Canaras Quiz)

PARC
Mark Weiser - Computer for 21st century
OS, garbage collection, user-interface design

Coined term in 1988
ubiquitious computing

- No one will notice their presence. they will be ubiquitous.
their → hardware - software elements

Ubiquitous computers.
→ location and scale — should know
than no need of artifical intelligence.

Tabs → smallest example of embodied virtuality.

no artificial intelligence needed
merely computers embedded in everyday devices. (world)

Tab → size less than or equal to sheet of paper like ID card etc

Pad bet^n sheet of paper to laptop or palmtop size.

Board like bulletin/notice board.

Interaction important

Ubiquitous computing

Cheap, low power comp with display

software → micro-kernel.

problem networking. → single device for an

Machines that fit the human environment

2nd link

Calm technology

- smooth transition bet^r periphery & center
  - better view of periphery / locatedness

(1) extension of

eg. Glass windows - locatedness
gives

individual must be in charge of moving things from center to periphery and vice-versa.

(2) Multicast

(3) Dangling string

Wearable computing

4 layers

(1) mobile phone
(2) Carry many peripherals
(3) sensors, sensing healthcare
(4) ... generation
Intra communication

4th Paper (About what next)

- Ubiq. comp. taking advtgs. of existing infrastructure and using it for new purposes.
- Cyberguide, RADAR - first indoor position tracking application
  air points, signal strength measurements etc.

"Your Noise is my signal"

truly multidisciplinary interactive field
focused to solve real-life problems

2D programming environment → 3D application
task of ubiquitous app. designers

Mobile phones - single source of information
but as per ubiquitous computing, it should be used by anyone (notion of cloud computing etc.)

Challenges:

① Should notify other domains Imp. of ubiquitous computing related research in their domain.

Optional Reading

CS issues in ubiquitous computing

Computer should not be centre of attraction or source of attention.

Virtual reality - extremely useful in scientific visualization and entertainment.
but has 2 flaws:
① not useful in day-to-day life (physical world)
② not cheap & cannot replace computers.

Tab – Notesized
Pad – notebook sized.
Board – large size (maybe bulletin board)

flexible hardware & software components

Power $P = C V^2 F$
reduce voltage to low power, so increase area

wireless use of near-field radios.
very less problems of multipath & interference.

pens second layer of IP target actual address, if person is on a visit

---

30/08/2016 Ubiquitous computing
→ Pervasive computing
→ 3rd Generation

Mainframe Era
- Scientific calculations
- Data processing

30's

PC Era
- Spreadsheet
- Document Processing

60's

UC Era
Inch, foot, Yard Devices
Calendar, *Communication, Location services

90's

People to Device Ratio
Many – 1
1 – 1
1 – Many

Steve Mann

XEROX PARC

- mouse, ethernet, 3D printing, ethernet → discovered here.

Marc Weiser

Tabs - connect wirelessly to basestation - 10kbps infrared. (track location)
Pads - 250kbps (use like a scratch paper)
Boards - large displays, pen input

1) Cheap
2) Software
3) Communication link

Calm Technology

Olivetti
- Indoor location system (active research)

Active Badge Location system

Context-aware computing

ClassRoom 2000 → Living Laboratories
Awarehome
Wearable computing

Place lab → Intel
Sense Cam → Microsoft

Smart-Its
hardware - flexible → can be embedded into many devices.

1. Reading
2. Python + Scipy/Numpy + Scikit-learn (Anaconda package)

1st Sep. 2016

Activity Recognition Pipeline

Sensor Data → Analysis (cleaning of signal) → Model

- Two Pre-dominant Sensing Approaches
  1. Environmental
  2. On-Body sensors.

Sensor Data Acquisition

Wearable Sensors (Accelerometer, GPS, → Integration Devices (Cellphone or Laptop or PDA) → Communication → Storage and Interference

Aggregation
Storage (Local vs Server) (helps to monitor if data is coming)
Visualization

Training Data $\xrightarrow{1}$ Preprocessing segmentation Frame, feature extraction $\xrightarrow{2}$ Train model $\xrightarrow{3}$ Model

Real Data $\xrightarrow{4}$ Preprocessing $\xrightarrow{5}$ Model

we don't train it 2nd time.

1

Preprocessing

Synchronisation
Validate sensing specs
Remove undesirable artifacts
Downsampling

Encoding
Missing values
unit conversion.
Quantization

Scaling to Range
Mean removal
Normalization

Smoothing
Moving average
exponential smoothing
Low Pass filter

Trend
Cyclic
Irregular fluctuations

You can use segmentation (Classifier) to remove unnecessary parts.

frame duration = 1/2 (actual desired)
Overlapping frames
Lot of heuristics involved.

2 Frame & feature extraction

$\begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$ feature vector Mean, variance, Kurtosis, skewness, RMS

* feature selection
  - Signal-based
    Statistical
    freq domain: MFCC ??, DCT, FFT
  - Physical model
    Limb trajectories
  - Extract from data
    Clustering, PCA, LDA

Classification

1. Decision trees
2. KNN,
3. SVM
4. Bayesian networks
5. Graphical models (HMM, CRF)
6. Symbolic Representations (Vector Space Model)
7. Deep Nets

Multimodal Input

Inertial → PSFF
Audio → PSFF
Location
Image → PSFF

→ (+) → Train model → model

OR

—11— → —11→ train model → Model
train model → Model

→ (+)

This approach is more preferred.

Performance evaluation

Accuracy
Precision → false positive
Recall → False negative

$$\text{f-measure} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Confusion matrix

Actual activity

Predicted activity

Domingos Paper

Things about Machine Learning

- Learning → Representation + Evaluation + Optimization
- Generalization counts
- Data alone is not enough / knowledge required
- Overfitting has many faces → bias & variance
- Intuition fails in high dimensions
- Theoretical guarantees are not what they seem
- Feature engineering is the key - right combination of features
- More data beats the clever algorithm
- Learn many models, not just one
- Simplicity $\not\Rightarrow$ Accuracy
- Representable $\not\Rightarrow$ Learnable
- Correlation $\not\Rightarrow$ Causation

ML-Mitchell

Computer Vision
Robot control
Speech Recognition
Bio-surveillance

ML = Comp science + statistics

Recent Questions

1) Unlabel data in supervised learning?
2) Transfer knowledge from one task to other?
3) Relation betⁿ diff. learning algorithms?
4) effective way to collect own training data?
5) Privacy vs. data mining?

Long term questions

1) ML programming language design?
2) Build never ~~[illegible]~~ ending learners?
3) ML ↔ human learning?
4) comp. perception merge with machine learning
5) Ethical questions: privacy

Smola Book

1) Naive Beyes
2) Nearest neighbour
3) Simple classifier
4) Perceptron 5) K-means

07/09/2016

$\theta = (X^T X)^{-1} X^T y$

- Activity Recognition Challenges

  Intra-class variability $\Rightarrow$ Diff. styles of eating food

  Inter-class similarity $\Rightarrow$ Eating and talking - mouth movement

  NULL class problem $\Rightarrow$ How to classify everything else apart from desired feature.

  Class Imbalance $\Rightarrow$ Different activities have diff. dataset.

  Ground truth annotation $\Rightarrow$

  Extracting Qualitative information $\Rightarrow$ in contrast to binary classification.

  Variability in sensor

  Operating requirements

  Sensor fusion

- Machine Learning (ML)

  field that gives computers ability to learn without being explicitly programmed.

  Types of ML:

  Information based: Supervised, Reinforcement, Unsupervised.

  Algorithms:

  Regression-based, instance based.

  Linear model

  $h(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2$ (try to minimize the error)

  $J(\theta) = \frac{1}{2} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2 \longrightarrow$ cost function

  $$\boxed{\theta = (X^T X)^{-1} X^T \vec{y}}$$

other approach: Least square cost function.

Decision tree

Each hypothesis is based on Research tree.

- Discriminative vs Generative

  → Hypothesis defines decision boundary between classes.

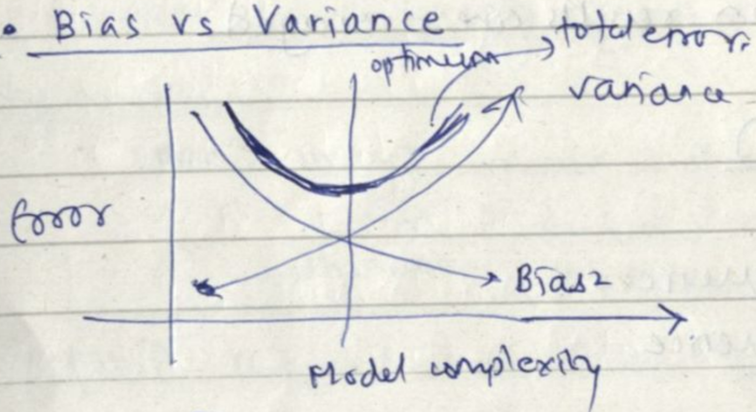
  Build probabilistic model for each class and match against each model. e.g naive Bayesian

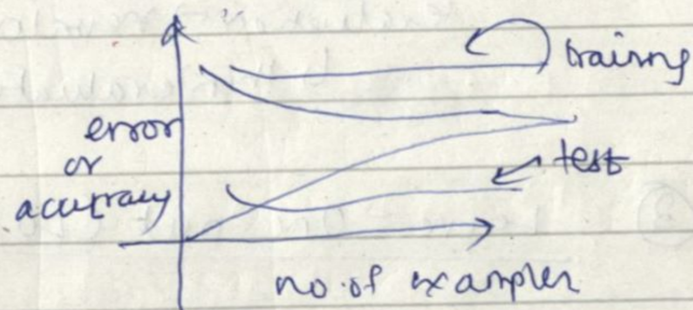
- Underfitting vs Overfitting

  Underfitting → Model not complex enough to fit the training data.

  Overfitting → Model too complex, learns the noise in the data.

- Bias vs Variance





no change in error no matter how much examples we take, threshold reached.

Need to decide between underfitting and overfitting.

Team → 20 Sept.

Proposal → 27 Sept.

Progress Report → 25 Oct.

Final report → 29 Nov

Presentation → 29 Nov

Detect stress from gestures

Activity recognition with physiological signals

AR models from media.

Swimming style

cooking gestures

dog activities

text & drive identification

Handwashing - wrist sensors

8th September 2016

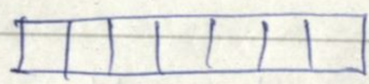
- Classifier Performance Evaluation

① Split Train / Test

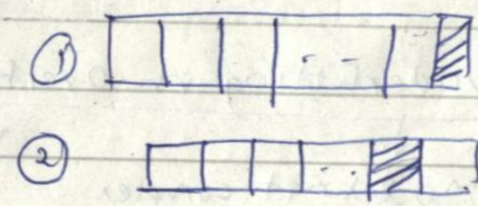
60% → Training Data
40% → Testing Data

Usually, we have validation data set reserved for tuning parameters after training & before testing

② Cross-validation



K-folds equal



K cases

Training → (K-1) folds
Evaluation → remaining fold
↳ then evaluation results are averaged.

③ Leave-One-out (LOO)

Training → (n-1) sequences
test → last sequence.

④ Leave one participant out (LOPO)

↳ good result → more generalized classifier
↳ very effective for human activity recognition evaluation.

Predict wine quality
diff. classifiers
using only
some features

Datasets

Train classifiers

Readings (Used bluetooth module for proximity detection of mobile)

Dey paper - Getting closer ; Empirical investigation of the proximity of the users to their smart phones

- based on previous paper by Patel
- 28 users → within arm distance - 53% of our time
4 weeks within (arm + room) distance - 88% —"—

Limitations

1) Loss of data because their application got killed by task killer app. (about 53%.)

2) time span of 4 weeks may not be sufficient to generalize.

My opinion:

~~[illegible]~~ 1) maybe we can include more subsets for distance.
1-2m, within arm, 2-4m, just nearby, 4-6m & greater → somewhere in the room (large distance)

Advtgs 1) Helpful for application developers, researchers to be aware of these survey results while designing the application.

2) Studying effects of proximity & relating it to healthcare (signals coming out from phone are harmful)

3) Models like ~~Random forest model~~ & decision tree classifier using ID algorithm

4)

Sensus : Cross-Platform, Mobile crowdsensing (MCS)

<table>
<tr><td>

1. Protocol
2. User-friendly
3. Can collect sensor data in background
4. AWS, comptabile with MATLAB, R

</td><td>

Limitation

1. iOS cannot probe in the background, requires notification to be sent which will trigger the event in turn.

</td></tr>
</table>

Kidd : Aware Home : A Living Laboratory for ubiquitous computing research

1. Smart floor , Ground reaction force profile to identify / track the person.
   using force-sensitive load tiles.

2. Finding lost objects (FLO → frequently lost objects) such as keys, glasses, wallets

RF tags on each object & indoor positioning system, LCD panel to track these objects.

→ Support for the eldery
→ Social issue evaluation

Future challenges

1. Qualitative understanding of everyday life of humans.
   → how they ~~access~~ find lost objects
   → which objects are lost frequently
   → how other members help them to identify & use these things to train their FLO system to be more intelligent.

20/09/2016

ABCs of Human Subject Research

Why

History

Protocols

Data → ISFF → Train Model → Model.

Human Research important

Why?

1. 600 people recruited for tracking natural progression of untreated syphilis. (Study 1932-1972)

   they offered free meals, told this is not harmful disease. but kids, wives affected

2. Milgram experiment (1961)

   ↳ Obedience experiments

   → administering electric shocks

   Teacher → Participant

   Learner → Actor. Teacher asked delivering shocks despite request of learner to stop

   Highly psychologically disturbing

3. Radioactive material to pregnant women,

4. many other

Belmont Report

- Respect for persons
- Do not harm
- Justice

IRB - Institute research board for reviewing the research & practices followed.

Consent form must be signed by all the participants.

Two classes of studies

① Lab Studies

Controlled settings

Easy to obtain ground-truth

Better internal validity

② field studies

Non-controlled settings

Hard to obtain ground-truth

Better external validity (more generalized)

ResearchKit → iOS framework to conduct medical research.

ResearchStack → To build research study apps for android.

22/09/2016

Environmental vs On-Body

1) camera → issue of privacy and other processing difficulties like computer vision.

2) simple sensors in home

↳ processing not difficult, but maintenance is a issue.

3) Kinect → Depth Cameras

4) LIDAR → becomes difficult to classify as on-body or environmental.

Specialized vs Commodity

Specialized Sensors

↳ for detecting eating etc.

Advtgs. vs Disadvtgs.

Advtgs: more accuracy.

Commodity

Assisted GPS

Touch ID,

Wifi/Bluetooth

Accelerometer → Acceleration

$F = ma$

$a = \frac{d\vec{v}}{dt}$

Gyroscope

angular velocity.

MEMS Gyroscope

mass shifts & then angular velocity is calculated.
(rotation)

→ Gyro-wheel car

Barometer: Air pressure, ~~Alt~~ Altitude & relative elevation.

GPS: 24 satellites, 4 per plane, 6 orbital planes

4 satellites required at any time

3 → co-ordinates

1 → offset time from standard time.

Assistive GPS

Instead of communicating with the satellite directly, they reach to cell tower.

- helps to acquire faster.

GLONASS

Global Navigation Satellite system.
(Russian-space based satellite navigation system)

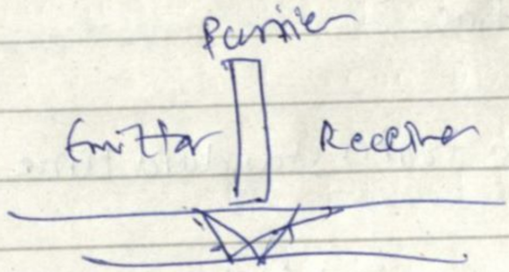
Accelerometer + Gyro
Orientation changes
Physical activity tracking
Games

Ambient light
Adjust display brightness
Battery conservation

Proximity — Turn off display
Barometer — weather forecast
Step count (Chairs)

for battery
co-processors are there, if not used, then they are turned off.

$\boxed{\text{Optical HRM}} \rightarrow$ heart-rate monitor
↓
(PPG) → amount of light scattered by blood flow.



Optical emitter
DSP
Accelerometer
Data Analysis
↓
Multiple wavelengths based on skin morphology

BSX → Austin company
↓
Optical Lactate threshold measurement → balance both aerobic vs anaerobic exercise.

Skin & In Body sensors

↳ Ingestible sensor

Galvanic skin response

↳ Skin resistance. (sweating)

27th September, 2016

, Collecting Sensor Data

*. CrMotionManager → Class for capturing data on iphone.

updateInterval = 2' ⇒ 200 ms.

· Photos:

AVCaptureInput

* Audio Recording -

AudioSession in Iphone.

RR change → heartrate monitor in MicrosoftBand.

A practical approach for Recognizing eating moments with wrist-mounted inertial sensing

- Ubicomp 2015 (actual 21) [13 male] Age 20-43
- Semi-controlled lab - 20 participants

		Ffactor	Precision	Recall
Right-handed	7 participants - 1 day	76.1%	66.7%	88.8%
	(38 year male) 1 participant - 31 days	71.3%	65.2%	78.6%

- Practical, automated system for food-intake monitoring (In-the-wild) study

Earlier Work

- devices like neck collars for swallow detection
- microphones inside the ear to detect chewing

Goal:

To distinguish betn eating and non-eating activities.
Eating moment detection using wrist mounted Pebble watch & getting accelerometer data (3-axis) (25 Hz)

Motivation

- Useful for researchers in healthcare field, dietician.,
- to moniter obesity and diabetes,
- dietary $\boxed{\text{self-monitoring not useful}}$ (wanes over time) problem of biases and memory recollection.
- Semi-automated food Journaling
  - Google class to capture photo on eating activity detection.

Related Work

Acoustic sensing - Piezoelectric gauge, Bodyscope, head mounted microphone, wearable camera

Wearable Cameras
- Problem of privacy and image analysis.
- Prof Thomaz → platemate to detect nutritional information from food photographs.

Inertial Sensing
- 5 inertial sensors - but complicated
  - it distinguishes only between eating and drinking and not other activities

Advantages of described method
- easy to use
- more generalised, evaluated at naturalistic conditions
- Public dataset available too further exploration

Pilot Study - helped to understand non-eating activity types, smartphone weight causing issues

other similar activities considered : walking, phone call, brush teeth, comb hair
- Added 3 seconds offset to provide buffer/compensation for diff. eating styles

battery-life ↑

Participants were asked to wear a camera for ground truth calculation - 60 seconds

Pipeline
Preprocessing - EMA filter
feature extraction - Mean, variance, kurtosis, skew, RMS
Intake Gesture identification (classification) → Random forest - better handling of non-linear data
Eating moment detection
↓
temporal density
DBSCAN clustering algorithm

frame extraction - sliding window approach with 50% overlap
frame size - 6 seconds on average

Random-forest with leave one participant out (LOPO) cross-validation, 1/3rd of data for testing.

DBSCAN → min gestures, distance measure

detect diff in

- eating with knife & fork
- eating with hands
- eating with fork or knife only.

Challenges / Limitations

- similar activities, hair comb, phone call, teeth brushing
- wiping face with napkin, scratching head
- most difficult was Chat.
- Intra-class diversity - diff-habits
- Instrumentation - misclassification, hand, knife etc
- Ecological validity - wearing watch in another hand
- Battery performance (photo, data sync on phone from watch, upload to server - only 5hrs 42 minutes)

future work

- Gesture detection using Dynamic time warping (DTW)
- new features like GPS & (more modalities.)
- more personalization.

Activity Recognition from User-annotated Acceleration Data

– Ling Bao – 2004

- 5 biaxial accelerometers – wire-free
- 20 participants
- Mean energy, entropy and correlation calculated.
- Best perfomance with decision tree – 84% accuracy.
- 2 accelerometers – reduction in accuracy drastically.

Recent work

- only under controlled lab conditions,
- use limited dataset and detect/recognize limited activities (7-8 like standing, sitting etc.)

- Only little work to validate the idea of acceleration under real-time circumstances. . .
- most of the data collected from researches & hand-annotated by themselves, so not useful or realistic.

- Permits users to train algo. if they label the data.

- right hip, dominant wrist, non-dominant thigh, non-dominant upper arm, dominant ankle

- Accelerometer – ADXL 210E (Analog) 2 axis
  ±10G – 2% tolerance 76.25Hz,
  90 minute data collection, hoarder boards.
- no extra sync. betn hoarder clocks
  belt synced to watch clock – offset 1-3 seconds.
- shaking was done (to measure time-shifting)
- T-mobile sidekick phone pouch used as a case.

Activity labels
- 20 activities
walking, brushing teeth, bicycling, sitting, eating

• semi-naturalistic, User driven data collection
- monitor laundry folding which does not occur on regular basis.

- given worksheet containing obstacles (obstacle course)
users can work at their own speed, start & stop time recorded, relevant notes of that activity.

• Specific activity Data collection
- more focused like describe walking, scrubbing
- still more freedom, can be done outside laboratory
- No researcher involved in data collection.

- feature computation
512 sample windows - 256 overlapping
76.25 Hz sampling frequency
Mean, energy, entropy, correlation - feature vectors.
fast computation of FFT.
DC - mean acceleration
energy - $(FFT)^2$ Sum.
entropy to distinguish betn same energy activities.
correlatn betn 2 axes - improved multiple body parts movement detection.

Evaluation

- Posters, mails across campus for recruitment.
  (13 males, 7 females. (Age 17-48)
  6 participants skipped through some obstacles in obstacle course.
  8 — 11 ——— 2nd session some tasks.

- Weka ML Algorithm toolkit - Decision Tree, IBL, naive Bayes tried.
  ↳ Best

  2 protocols
  1) Train - sequence data of 20 participants
     test - obstacle course data.
  2) Train - sequence f obstacle course data - all except one
     test - on one subject's data.

  2nd Best - Nearest neighbor

  stretching missclassified as folding laundry.
  riding elevator ⇔ riding escalator

- user specific training data improves accuracy of the classifier (collected data from companion & tested this)
  - shows that pre-trained systems are useful for real-world applications & offers great benefit.

- Accelerometer on thigh - most influential for recognition.
  then dominant wrist than non-dominant arm.
  then 2nd best - accln of hip.
  carried out by leave - one - accelerometer - in testing.

wrist - upper body (TV, sitting) recognition.
ankle - lower body

teare two accelerometers in testing
best is thigh & wrist
but still 5% less than effective of all accelerometers
& also hip & wrist.

Analysis
- User-specific training - not necessary for getting 80% accuracy

- here 84.26% prediction/ recognition rate - good deal.
C-4.5 classifier
- bicycling - hip acclⁿ energy
running - high entropy in hip correlation and higher hip arm movement correlation.

- Temporal info can be used.
standing still vs riding on elevator
↳ takes few seconds.

- user-specific training required to use date & time data.

Limitations - decision trees slow to train - but quick to run.
- does not recognize activity style like slow walk, fast walk etc.
- exact pace of walking??
e.g. use of heart-rate data, GPS & other modalities to improve accuracy.

4|10|2016 **Inertial Sensing**

Commodity vs custom sensors

Commodity → lot of features, well engineered

custom sensors → designed by own.

/ Time series Analysis → uses ECG.

Instance → unit of analysis.

1) Overfeeding
2) Lot of noise

} if we directly feed x (input) from sensor to recognize activity.

↓

Preprocessing done (feature extraction)

Time Domain

1) remove noise (Moving average filter)

↓ take window, Find average, slide window.

Disadvantages:

1) This filter has delayed peaks.

So, we use.

(2) Exponentially weighted moving average (EMA)

weight for samples.

* (3) Median filter → no lag, removes spike noise.

↓ Blurring

Image Reconstruction.

Frequency Domain filtering

(1) Fourier transform (LPF)

e.g walking on a treadmill

use mean removal, normalization

Frame & feature extraction

→ size of the window

→ overlap

→ Features: (Representational Learning)
ZCR (Zero crossing rate)

GOAL: Distinguish the frames that we want to model.

Low computation and storage cost

How to find!

11/09/2016 Location as Sensing modality

Time, change, Repetition

Location Information
- Absolute
- Relative
- Symbolic (area names)

Determine Location
- Client based
- Network-based
- Network-assisted

Network-based → infrastructure computes location e.g. Active Badge client can be lightweight

Client based → Client computes location No infrastructure required. GPS

Network-assisted → Combination of the two

Determine Location

- Proximity → wifi access point , stepping on pressure sensor
- Time of flight →
- Triangulation → Measure angle of arrival
- fingerprinting → pattern matching
  (temporal and spatial stability)
- Dead Reckoning → Kalman filter

13/10/2016

Computer Vision

- make computer learn from images, videos

high freq. information is tough.

- OCR (Optical character detection)

Face detection

smile detection

Object Recognition in Supermarkets

Vision based biometrics (Iris)

forensics

Object Recognition (Mobile Phones)

Google Glass

Shake, motion capture, Sports

Smart, Google cars   Robots (Industrial),

Medical Imaging

Why difficult?

1) Viewpoint variation
2) Scale
3) Illumination
4) Intra-class variation
5) Motion
6) Background clutter
7) Occlusion

Image

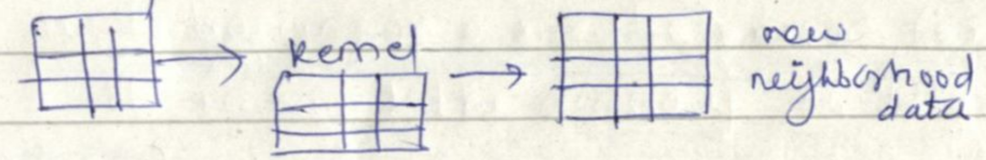
- Array of pixels (matrix representation)

  $0 \equiv$ black, $255 \equiv$ white

  $g(x,y) = f(-x,y) \Rightarrow$ rotate (flip) the image

Linear filtering

- Apply kernel



Sharpening (Accentuates image)

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

(remove neighborhood) (blur function)

- **Characterize edge** (sudden change portion)
  - calculate derivative.

$$\frac{\partial f}{\partial x} : \begin{bmatrix} & & \\ 1 & -1 & \\ & & \end{bmatrix}$$

- **Image Gradient**

$$\|\nabla f\| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$$

- Effects of noise

  ↓

  smooth it & then find edges.

Recognition

Visual objects: ~10000 to 30000.

Categorization:
- Indoor/Outdoor

Recognition
- all about modeling variability.

Allignment - do transformation to minimize residual.

Identify keypoint features: Peaks, corners, edges
Textured patches
Large contrast changes

SIFT, MOPS descriptors of features

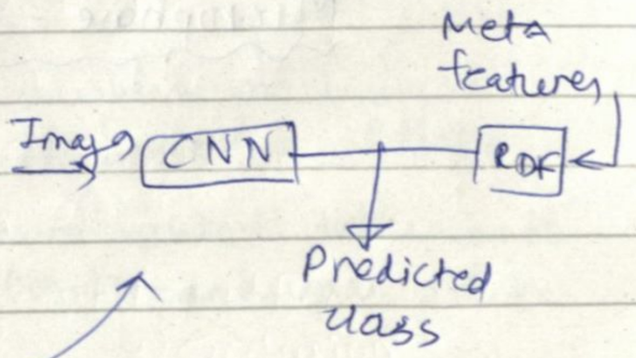
Bag of features outline
1. Extract features from variety of images
2. Learn "visual vocabulary". → use clustering to have codebook or vocabulary
3. Quantize these features
4. Represent images by frequencies of "visual words".

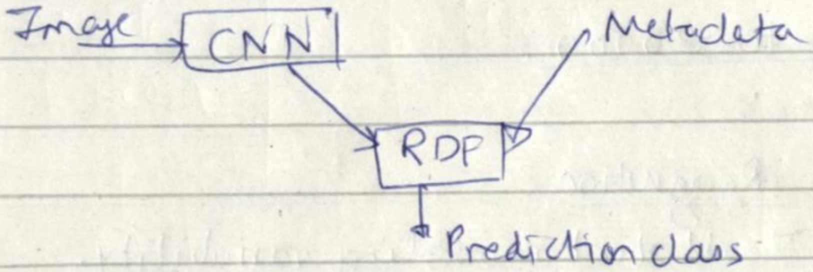
Weighing the words
a word that appears in all documents is not useful. We cannot use it for differentiation.

Paper

Convolutional neural network

- Classic Ensemble
- Late fusion ensemble.





83.07%

- Model did not generalize in general for all people

but,

with training on a specific person for 1 day and testing it improved the accuracy 76.92%. (↑ 65.44% when we got bad results.)

Computer Graphics: Models to Images
Photography: Images to images
Vision: Images to models.

Kristen Grauman

18/10/2016

Audio sensing

Outer ear
Middle ear → works as a transducer.
Inner ear

Microphone - Cheap hardware

Consumer Electronics

- Noise cancellation headphones
- Siri, Amazon Echo

Health

- Hearing Aid
- Spirosmart: To measure lung function on a mobile phone.

Types of Sounds

1) Noise
2) Natural Sounds (Animals)
3) Artifical Sounds (Machines, car)
4) Speech
5) Music

Lake Analogy

An auditory scene Analysis
- Source Separation (Cocktail party)
-

Sampling theorem (ADC)
- Nyquist

Audio Signal processing
Temporal Domain
Frequency Domain

Audacity Tool
Spectrogram (Image to audio sound)
Synthesizers

Audio features
- Pitch fundamental frequency
- ZCR (Zero crossing rate)
- Energy
- Spectral flatness (Noisy or harmonic)
- Spectral centroid (Average frequency of the signal)
  - " envelope descriptor
  rolloff freq. below which most signal exists.
  flux rate of change
  MFCC ⟷ audio spectrum which takes into account non-linear human pitch

Tools:
Yaafe
Librosa

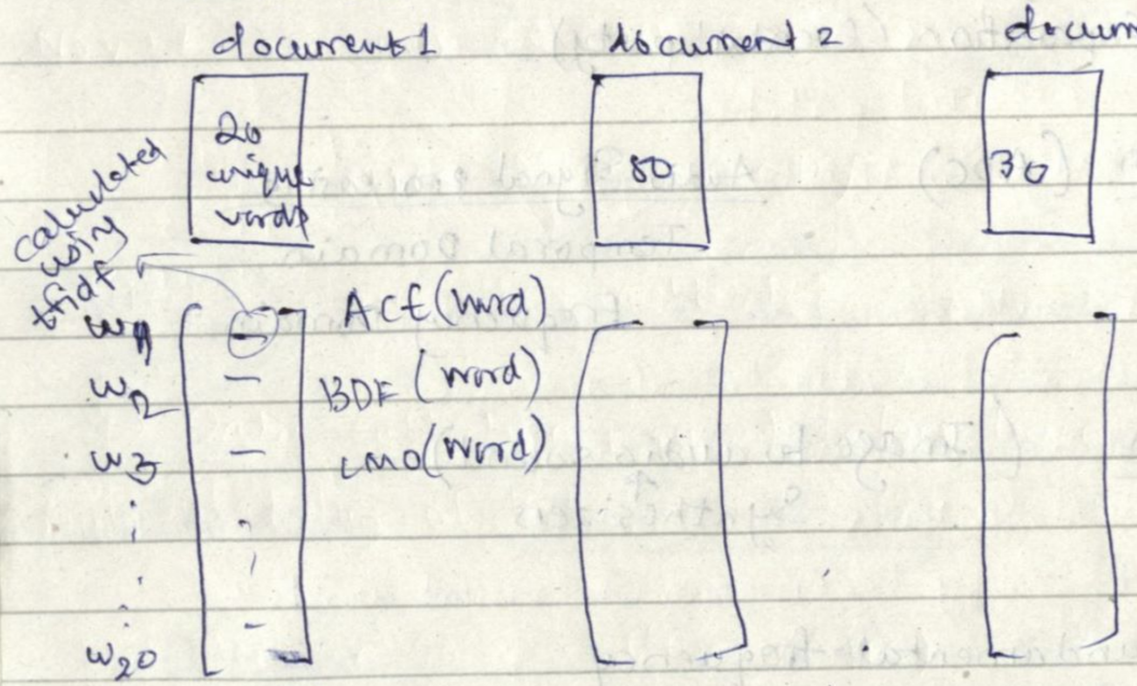
Panel of Experts

Matlab

HMM - commonly used for speech recognition.

Vector Space Model for Activity Recognition

- Rank documents



The angle of the query vector is used for ranking.

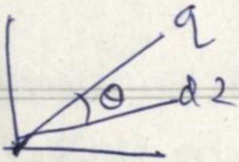
i.e. every vector will have certain angle with query vector.

- Documents and queries are represented as vectors

  if term occurs in document, its value in the vector is non-zero.

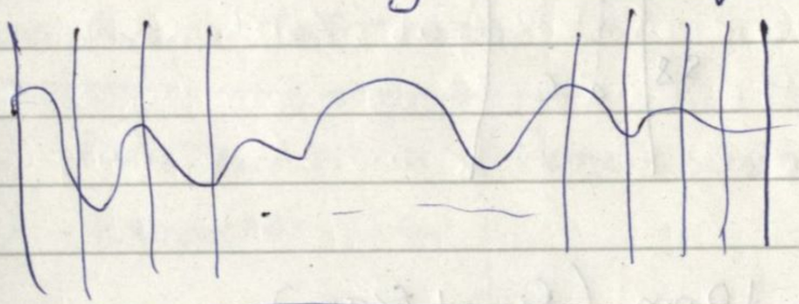
tf-idf
→ ~~[illegible]~~ term frequency - inverse document frequency.

$$w_{t,d} = tf_{t,d} \cdot \log \frac{|D|}{|\{d' \in D \mid t \in d'\}|}$$



$$\cos\theta = \frac{d2 \cdot q}{|d2|\,|q|}$$

n-gram technique



Similarity matching
Euclidian distance Metric

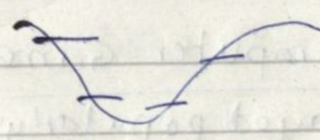
General Data Mining Algorithm

Make approximation & store it in memory.
but it should preserve the features.

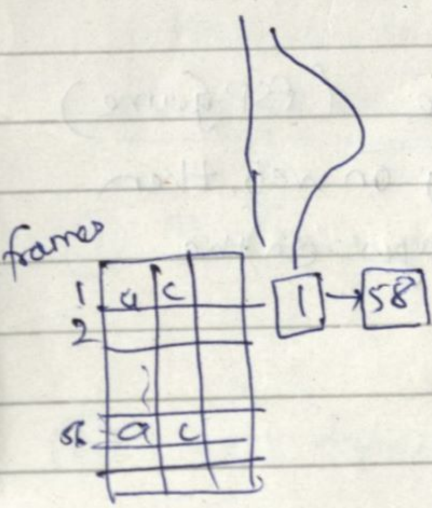
Lower bounding

- The distance after approximation should be less than or equal to euclidian distance.

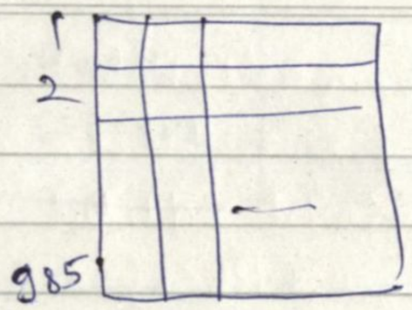
Symbolic distance (SAX)



allows pattern matching.



Symbolic representation for each frame.
for repetition form buckets.
& also form collision matrix for them.



1	a	c
2	b	c
⋮	⋮	
58	a	c

27/10/2016

Dec 8 ⟹ 7 to 10pm (final Exam)

Nov. 17 → Review session

Outsource ML machinery

Why explore other methods?

- → Data collection required before training the model
- → labeling of data.

So, Human computation (HC) ⟹ Luis Ahn.
↳ Integrating people into computational processes for problems which are hard for computers.

e.g. - Labeling images with Computer Game (ESP Game)
if the game is played popularly on web, then we will label almost all the images on the web within few months.

Recaptcha

Duolingo ⟹ Latest

Precision → false positives
Recall → false negatives

Amazon Mechanical Turks

Make money by working on HITs (Human intelligent tasks)

Case study by Prof Thomaz

1. Participant view (Remove sensitive images)
2. Researcher View

images fed to Amazon Turk HIT'S

1. One image per hit — expensive
2. All images per hit - not good
3. Group of images per hit based on time (1 hr) — Good (optimum)

(Decide based on voting from 3 people)

Calculate false P, False N, TP, TN based on the feedback got.

Master Workers — Amazon ranks - based on the prev. tasks completed by the people.

Precision low, lots of false positives
Recall low, lots of false negatives

⟹ VVIMP

1. This was due to lot of images in group. even human coder can miss one picture which is the only one to detect food/eating moment.
2. Meal type - not easily distinguishable — eating chocolate (snack or meal ??)

③ Not enough visual information.

But, this defeats the purpose of deployable (automated system as most of the tasks are done manually like filtering sensitive images because of privacy concerns.

Active learning – semi-supervised learning.

HMM → Probabilistic transition models.

Q. Vector space Model?? (what's the relation with n-gram technique).

1/10/2016 Environment sensors

Computer → can be used as a biometric sensor.
↳ detect changes in motor and cognitive function.
↳ Mouse events, login events tracking.

Drawbacks:
① Hard to install (Tens or hundreds of devices)
② Power often an issue (Battery or connected to outlet)
③ Hard to maintain (software update)
④ Collecting data (wired or wireless)
⑤ security (Explores a new vulnerability)

RF

6. Privacy (Lots of data collected about individuals)
7. Aesthetics (unpleasant to lookat, interfer with normal activities)

- Advantages

1. Nothing to wear
2. Hard to tamper with infrastructure.
3. Nothing to charge

3/11/2016

Infrastructure mediated sensing electrical, eg. air conditioning systems

- leverages existing home infrastructure.
- infrastructure activity as a proxy for human activity.

Who invented it?

- Shwetak Patel - one of the key inventors.

Advtgs:

1) simple sensing (only one device)

Disadvtgs

1) Lot of processing machine learning, data analysis required on single machine.
2) Data is highly position dependent.
3) Lot of complex software required.

Components:

1) Water
2) Electricity
3) Gaslines
4) HVAC
5) Network router
6) TV Cable

Applications developed:

1. Human activity, location
2. Measure human health
3. Track utility consumption
4. Behavior change.

08/11/2016 Applications of machine learning

HCI

Before that,

Hydrosense

↳ previous fixture identification

↳ Now activity recognition.

Now, Human Computer Interaction

Study of interaction betn humans & computers

Make this interaction natural, intitutive and seamless.

Doug Engelbert → Inventor of mouse.

Leap Motion Sensor

Body as an input interface (nails, hands, bio-sensing)

- used in watches.
- Pranav Mistry (video)

10/11/2016

Activity Recognition in Health

15/11/2016 **Privacy in Activity Recognition**

Fitness tracker → save money on health insurance
↳ privacy compromised.

Sound Shredding

- ACM Double Column format (Project Report - 15%)
  CH1 (Template available)

4 pages long report.

Include additional work

$\boxed{\text{Dec 4th}}$
@ 11pm

Final Exam

1) Activity Recognition Pipeline (Block diagram) why?
2) Data Collection
   - Different methods, settings, pros and cons
   - what are the challenges and why
     (lab, field, living laboratory)

   Ground Truth (very important term)

3) Machine Learning
   - Learning Algorithms, advantages and disadvantages
   - How to choose parameters.

   eg. Decision tree vs nearest neighbor.
   (when to choose one over another.)

   Issues like performance in wearable devices.

- what does it mean to have a good model?
  - How to evaluate a model (different techniques)
  - which metrics are often used?

(Person dependent vs independent)

(How to identify complexity?)
overfit vs underfit
(too complex) (too simple)

why Accuracy cannot be sufficient?
why Precision, recall?
confusion matrix

4) Sensing

Environmental vs On-body, pros and cons, challenges
Modalities and approaches for environmental sensing.
Extracting features, how to choose features (based on modalities)
- window size, overlap.
- Infrastructure-mediated sensing
- features useful for inertial sensing, freq. analysis
- which sensors to use when using Arduino.

Limitations of phone
Process of obtaining sensordata from phone (API's)
Hardware challenges in environment and wearable sensing

GPS, dead recokning, triangulation ← Techniques for determining location
Why is hard to work with images as sensor data?

Applications, HCI / Analysis of data, model

Privacy

Questions will be performed / framed like this.

10 Questions approx.

Emerging Topics

- Convolutional Neural Networks
  - Require lot of data
  - Used for image classification.
  - Also requires labeled images

that's why ~~the~~ is technique is not used so often, but nowadays this trend/use is increasing.

Scaling, Rotation problems are there.

Clarifai

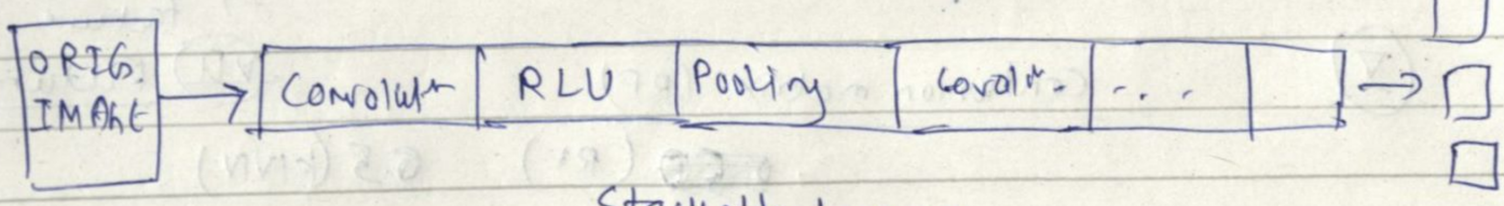
convolution works as a filter.

feature mini-image is compared with all the parts of original image.

Pooling, Rectified linear units (RLU)

values below 0 round them to zero.

ConvNet Architecture



Stackable layers.

Last layer - fully connected layers - to decide the result whether 'X' or 'O'.

Slides

I. Classifier

1) Labels, shuffled
2) frame, step. iteratively: 150 to 200 (5)
25 to 100 (5)
3) non-linearly seperable.
- SVM
- KNN $\boxed{k = 4}$
- RF.

II. Accuracy

- frame size = 90, step size = 80 RF
- Similar procedure, KNN
frame size = 105, overlap = 75

III. ~~The~~ results

- person-independant fashion.
- $\boxed{\text{SVM = 58\%}}$ (overhead) execution time for SVM much greater than other classifiers
- KNN = 72.28%
- 79%

IV. Confusion matrix (KNN)

- first column & 5th column.
RF

VI. Conclusion
1) RF
2) feature

VII. Future

V. Confusion matrix (RF)

~~0.56~~ (RF) 0.5 (KNN)
0.25

0.74 (KNN) 0.35 (