Case study on Automatic Facial Expression Recognition

SVMs In Action

Mitchel Benovoy

Outline

- The facial expression recognition problem
- The important of **pre-processing** your data
- SVMs for feature selection
- SVMs for classification
  - Multi-class approaches with SVMs
  - The black-art of tuning the parameters

The Facial Expression Recognition Problem

Anger

Fear

Happy

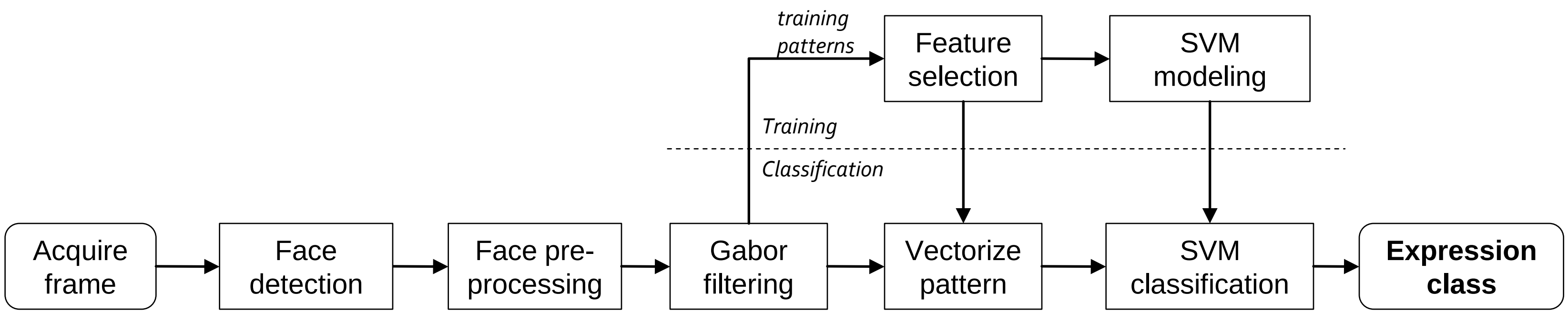
- Motivations:
  - Enhance human/robot-computer interaction
  - Design games for expressivity training (children with autism)

Constraints and Expectations

- Real-time operation
- Generalizes well
  - Gender
  - Race
  - Face type
  - Facial hair
- > 90% classification accuracy

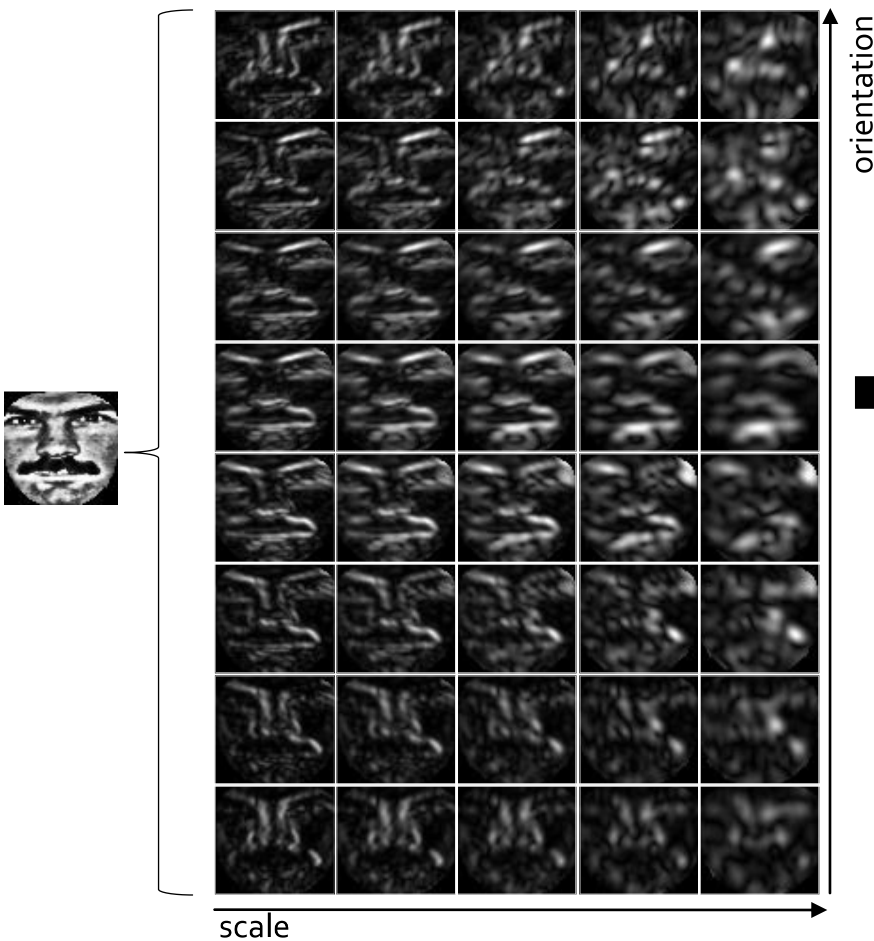
➡ **SVMs seem like the perfect candidate**

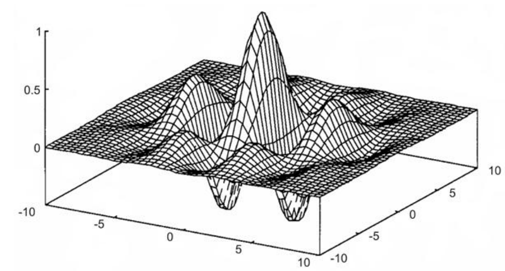
Process Overview



Gabor Filtering

- Exhibits optimal joint spatial locality and frequency information





Vectorized each sample to **1** x ***d***

255	246	67	78	120	236	201	...

- Downside to Gabor filtering:

Sample vectors are very high dimensionality:

For a **32** x **32** face image, ***d* = 40960**.

Data Matrix Formulation

- Given ***c*** classes and ***n*** total samples:
- Construct a label vector **Y** and a matrix **X** containing all the data:

$n \times 1$

Y =

1
1
2
2
3
3

$n \times d$

X =

255	45	97	197	…	205
247	56	84	200	…	222
236	78	99	245	…	128
201	24	120	54	…	145
254	65	101	97	…	135
222	74	134	213	…	20

Data Normalization for SVMs

- One of the most crucial steps when using SVMs
  - The optimization routines are sensitive to large numbers.
- Goal: map data to a standard range
- Two popular **methods** to normalize data:
  - Range normalization
  - z-Scoring
- Two popular **types** of data normalization:
  - Feature-wise
  - Sample-wise

Range Normalization

- Maps a series of values to a specific range
- E.g., [-1:1], [0:1] … [L:U]

V =							
255	246	67	78	120	236	201	147

$$V^* = \frac{V - \min(V)}{\max(V) - \min(V)} * (U - L) - 1$$

Here, [L:U] = [-1:1]

V* =							
1.0000	0.9043	-1.0000	-0.8830	-0.4362	0.7979	0.4255	-0.1489

z-Scoring

- Standardizes a series of values to zero-mean, unit-variance:

V =	255	246	67	78	120	236	201	147

⬇

$$V^* = \frac{V - mean(V)}{std(V)}$$

⬇

V* =	1.1352	1.0168	-1.3392	-1.1944	-0.6416	0.8851	0.4245	-0.2863

Feature-Wise Normalization

- Each column is a feature (pixel value in this case)

X =

255	45	97	197	205
247	56	84	200	222
236	78	99	245	128
201	24	120	54	145
254	65	101	97	135
222	74	134	213	20

X* =

1.00	-0.22222	-0.48	0.497382	0.831683
0.703704	0.185185	-1.00	0.528796	1.00
0.296296	1.00	-0.4	1.00	0.069307
-1.00	-1.00	0.44	-1.00	0.237624
0.962963	0.518519	-0.32	-0.54974	0.138614
-0.22222	0.851852	1.00	0.664921	-1.00

Sample-Wise Normalization

X =

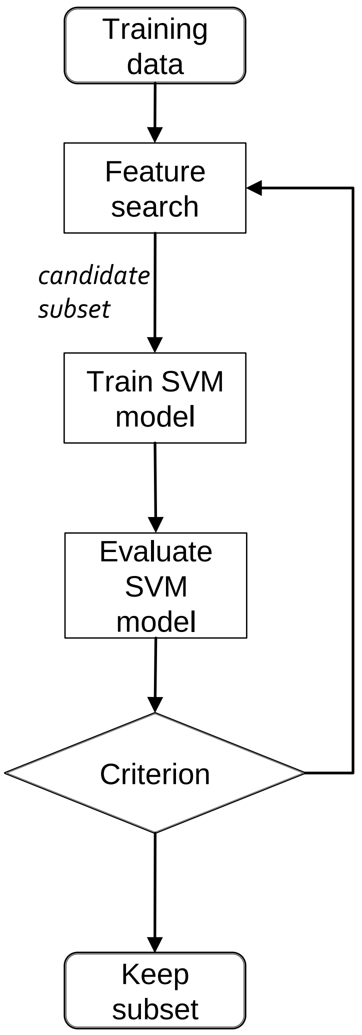
255	45	97	197	205
247	56	84	200	222
236	78	99	245	128
201	24	120	54	145
254	65	101	97	135
222	74	134	213	20

X* =

1	-1	-0.50476	0.447619	0.52381
1	-1	-0.70681	0.507853	0.73822
0.892216	-1	-0.7485	1	-0.4012
1	-1	0.084746	-0.66102	0.367232
1	-1	-0.61905	-0.66138	-0.25926
1	-0.46535	0.128713	0.910891	-1

SVM-Based Feature Selection

- Goals:
  - Reduce dimensionality
  - Find the salient features that discriminate between classes
  - Improves generalization
  - Removes redundant and irrelevant features

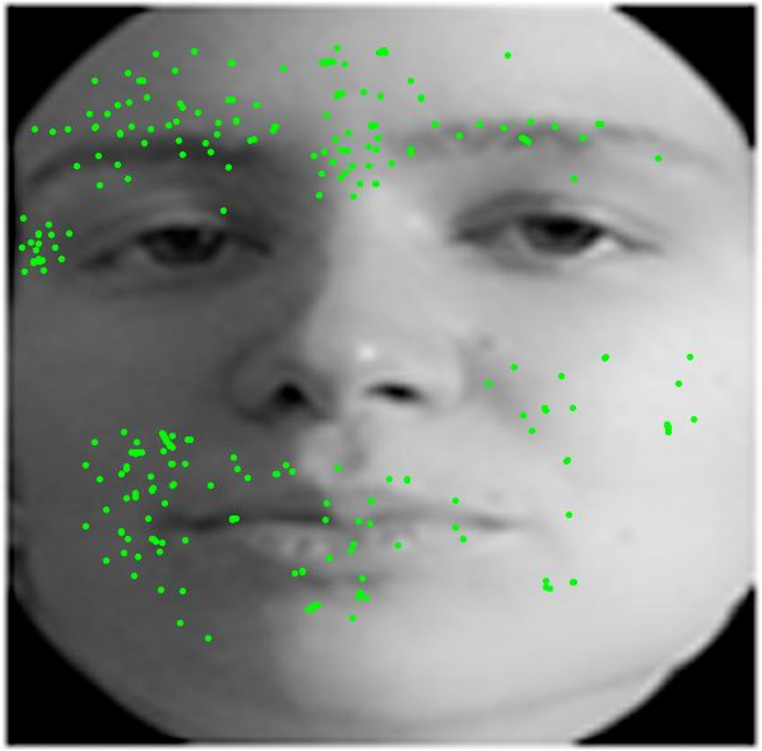


Salient Facial Expression Features

Feature selection is a great way to explore data.

Always check the selected features. Sometimes the searches get *stuck* or output garbage features.

SVM-based method



Correlation-based method

Multi-Class SVM Modeling

- One-vs-Rest
  - Train *c* classifiers:

Given 4 classes: {happy, sad, anger, neutral}

Train 4 binary SVM models:

1		-1
Happy	↔	sad, anger, neutral
Sad	↔	happy, anger, neutral
Anger	↔	happy, sad, neutral
Neutral	↔	happy, sad, anger

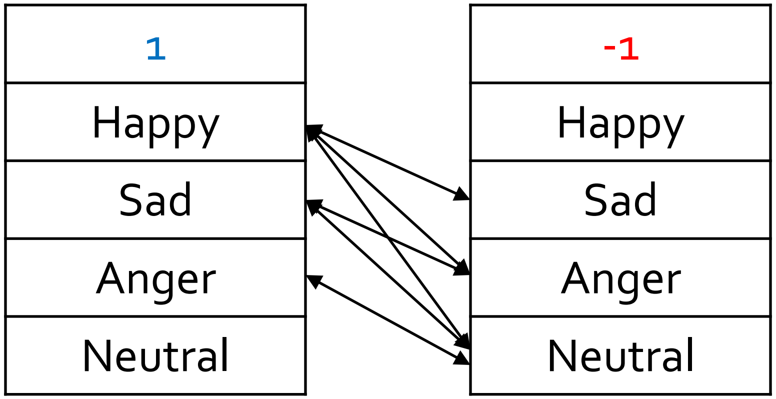
- During recognition, test each classifier and pick the positive class (1) with the largest margin (score).

Multi-Class SVM Modeling

- One-vs-One (Pairwise modeling)
  - Train ***c(c-1) / 2*** classifiers:

Given 4 classes: {happy, sad, anger, neutral}

Train 6 binary SVM models:



• There are more classifiers, but training time decreases as there are less data for each model.

• During recognition, pick the class that is selected by the most classifiers.

SVM Parameter Selection

- SVMs typically have 2 parameters:
  - Kernel-related variable (***g***)
  - Cost variable (***C***)

- Use grid-search approach
- Manual tuning

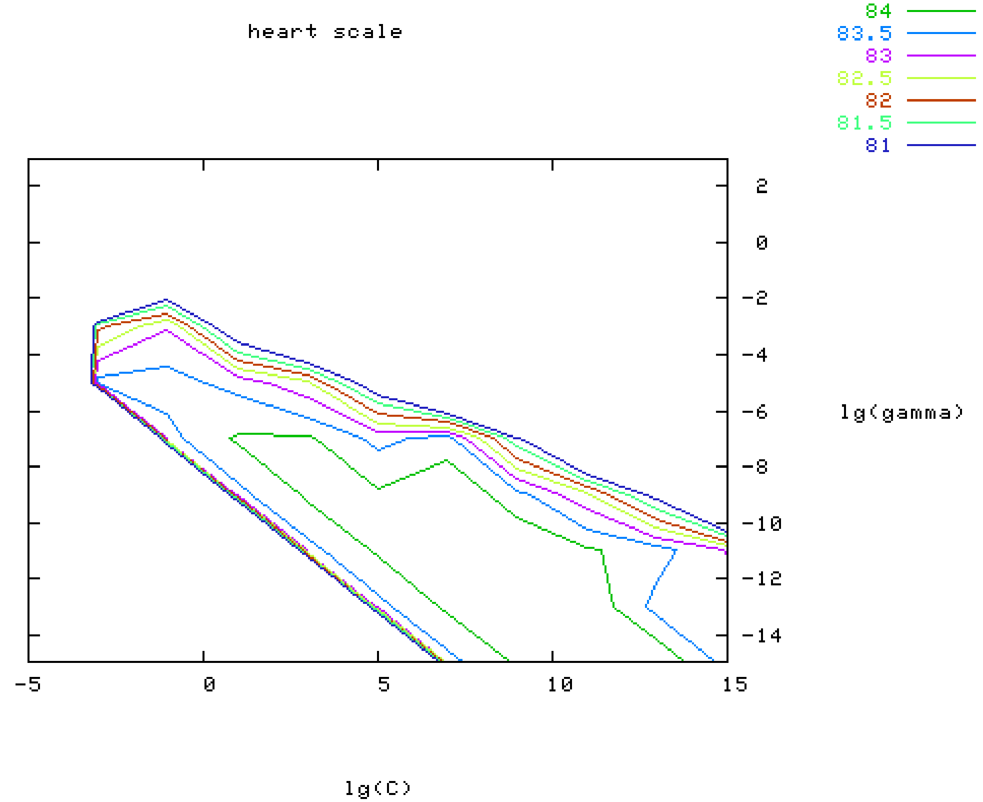


Image from http://www.csie.ntu.edu.tw/~cjlin/libsvm/

Performance

- 96% recognition rate on large, cross-cultural DB (Cohn-Kanade)
- 30+ fps
- Uses webcam as input
- Video…

Thank You

- Useful SVM links:
  - http://www.csie.ntu.edu.tw/~cjlin/libsvm/
  - http://nlp.stanford.edu/IR-book/
  - http://svmlight.joachims.org/
  - http://www.support-vector-machines.org/