Facial expression recognition using svm classifier on salient mic-macro patterns

Housam Babiker, Randy Goebel and Irene Cheng

Department of Computing, University of Alberta, Canada

Outline

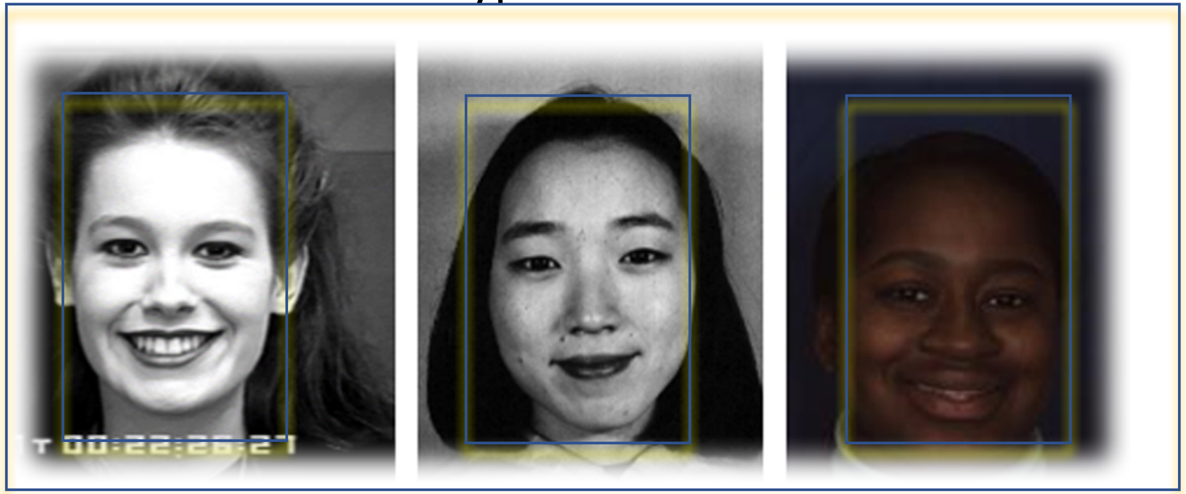
- Introduction
- Framework
- Mic-Macro patterns
- Experiments results
- Conclusions

Outline

- **Introduction**
- Framework
- Mic-Macro patterns
- Experiments results
- Conclusions

Introduction

- Facial expressions can be seen as the changes on the face due to the responses of our brain to social communication, emotions and intentions (*Huang et al*).
  - We need to be able to recognize all types of emotions.
  - To better understand the different types of the same emotions

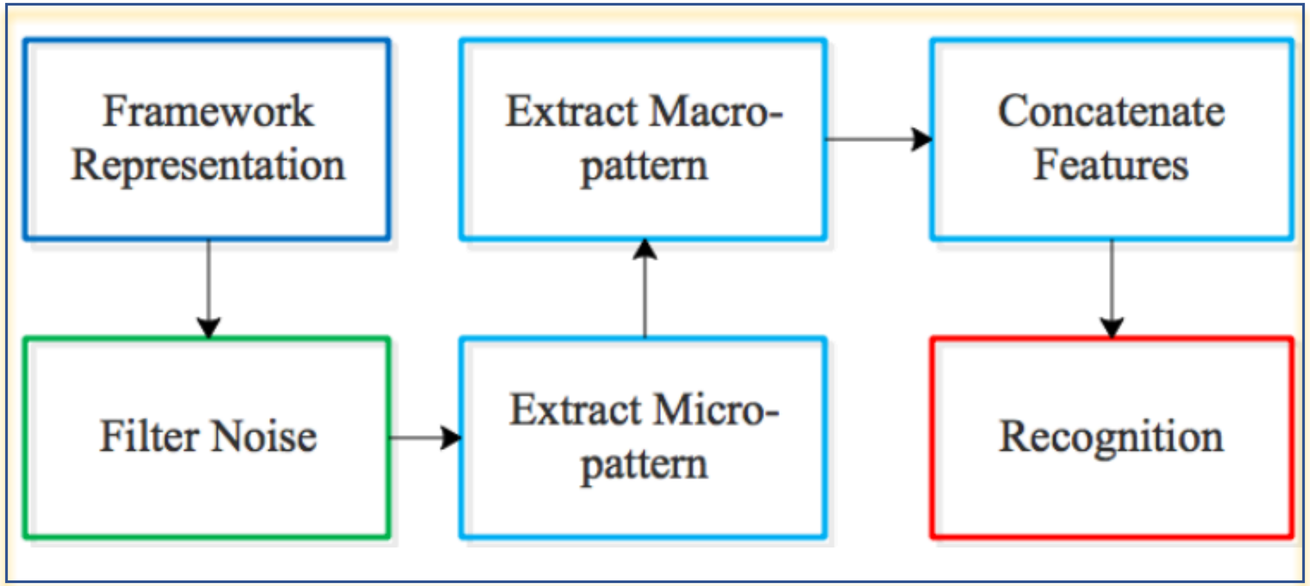


Expression of one person might change at different times and conditions.

Outline

- Introduction
- **Framework**
- Mic-Macro patterns
- Experiments results
- Conclusions

Framework


Framework Representation
Filter Noise
Extract Micro-pattern
Extract Macro-pattern
Concatenate Features
Recognition

Framework

- We introduce a dual-view perspective using two faces simulating a far-view and near-view.

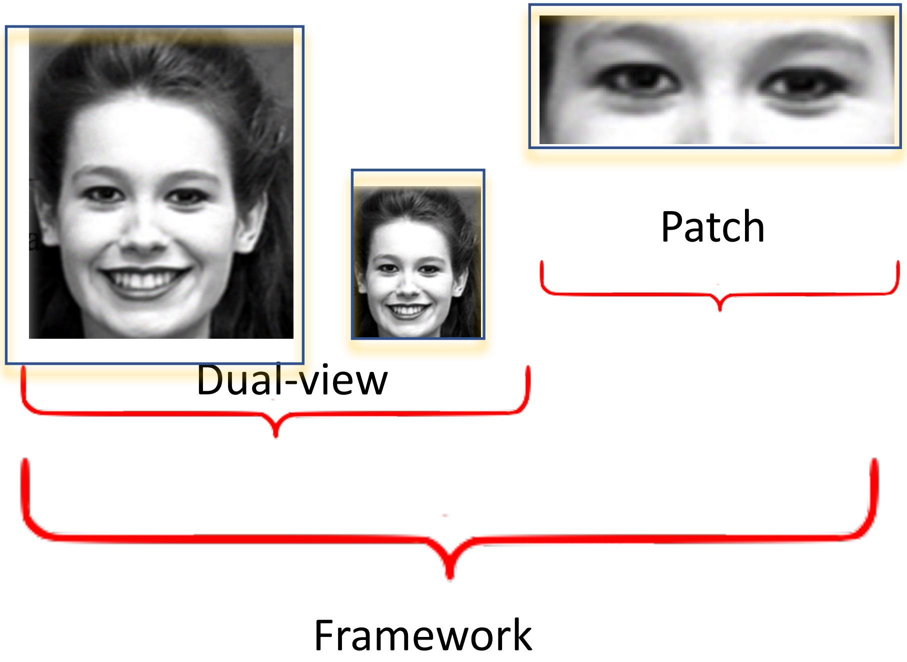
We made the choice to filter the noise as follows:

$$I^{'} = I - \mu_I$$

where I is the noisy patch and μ_f denotes the mean over the whole patch defined below:

$$\mu_I = \frac{1}{N-1}\sum_{n=0}^{N}(I_n)$$

where N is the number of pixels in the patch, and I_n denotes the n^{th} expression pixel.

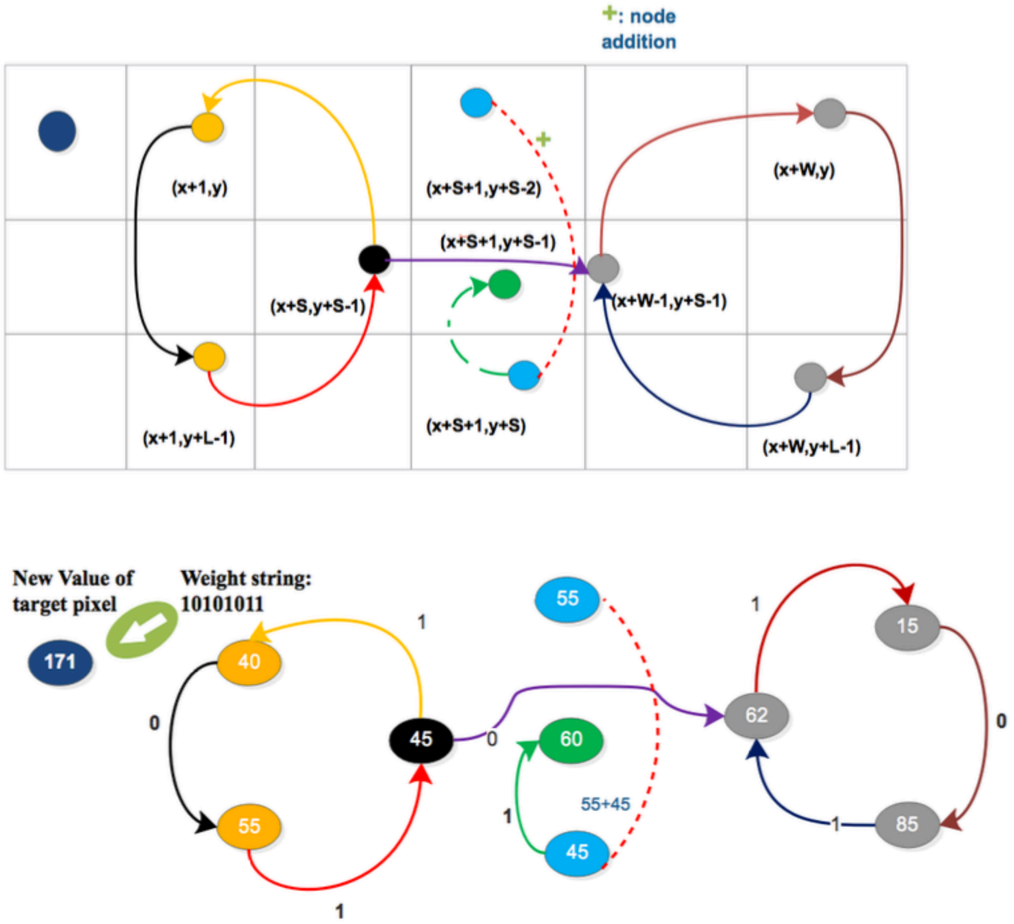


Chao, et al.," Facial expression recognition based on improved local binary pattern and class-regularized locality preserving projection." Signal Processing. '15

Outline

- Introduction
- Framework
- Mic-Macro Patterns
- Experiments results
- Conclusions

Macro pattern



Neighborhood

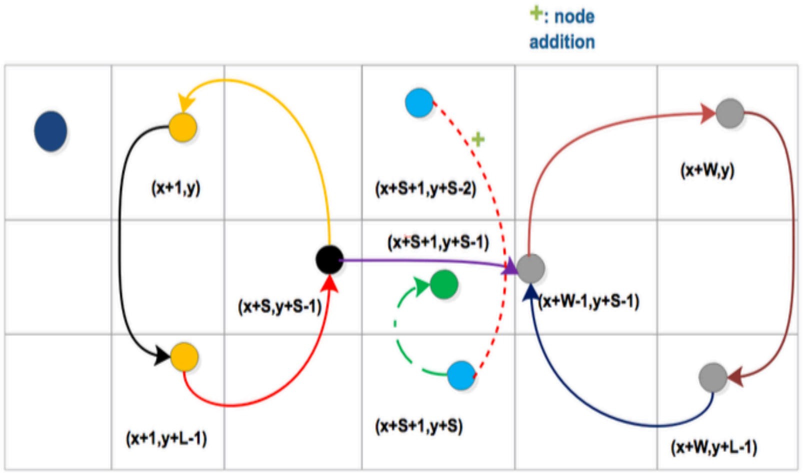
$$A = \sum_{i=x+1}^{x+W} \sum_{j=y}^{y+L-1} I'_{ij}$$

locations have the following dependency

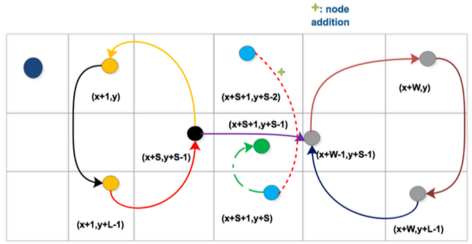
$$W = (4 * R) + 1, L = (2 * R) + 1, S = R + 1$$

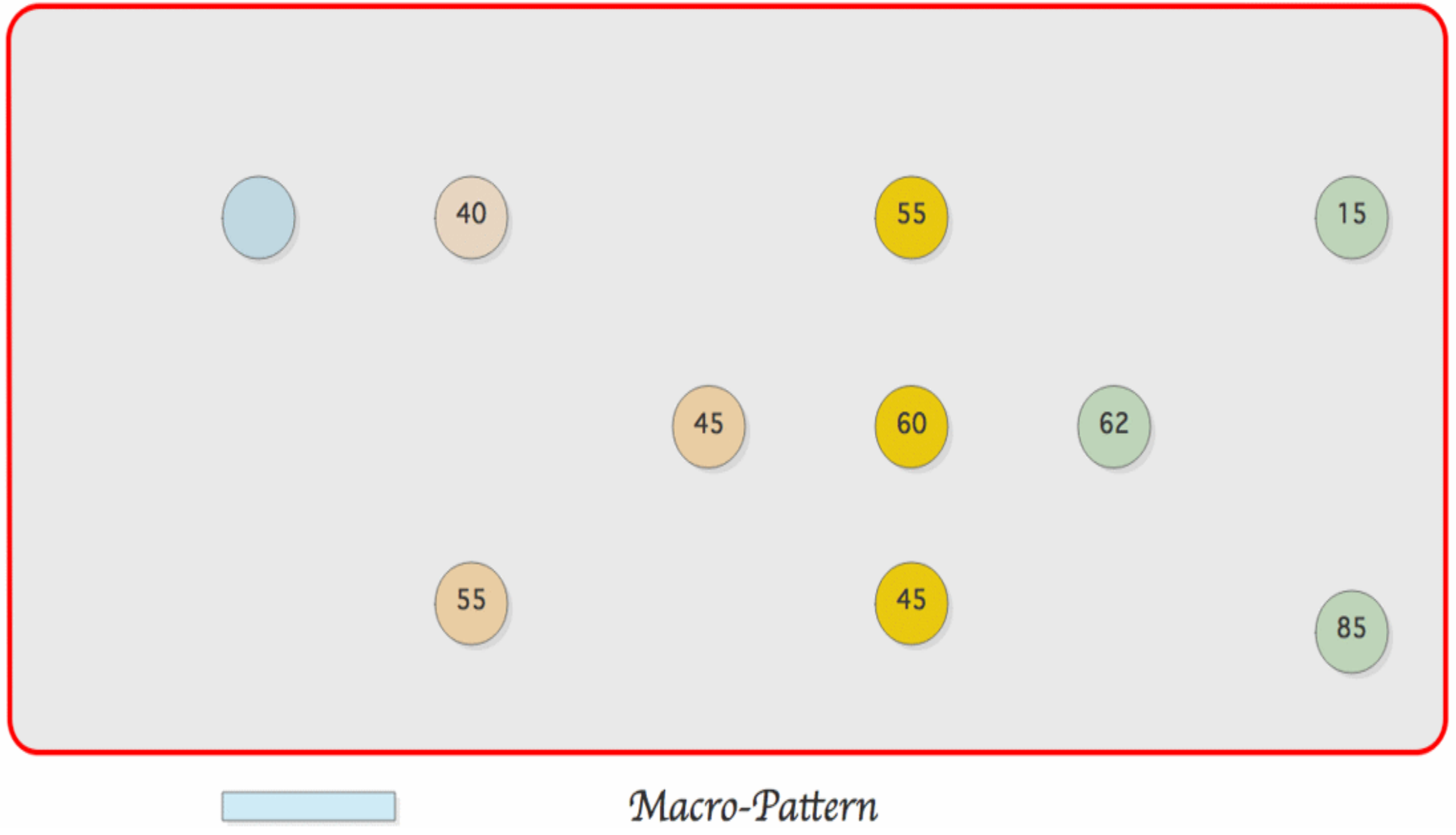
Macro pattern

- As the radius *R* increases, the number of non-graph nodes also increases, moving the starting black node further away from the target pixel.
- We realize that pixels close to the target pixel often *display similar features*, and pixels far away may display characteristics *irrelevant to the target*.
- Therefore, by increasing *R* gradually, we are able to collect neighbor-hood information.

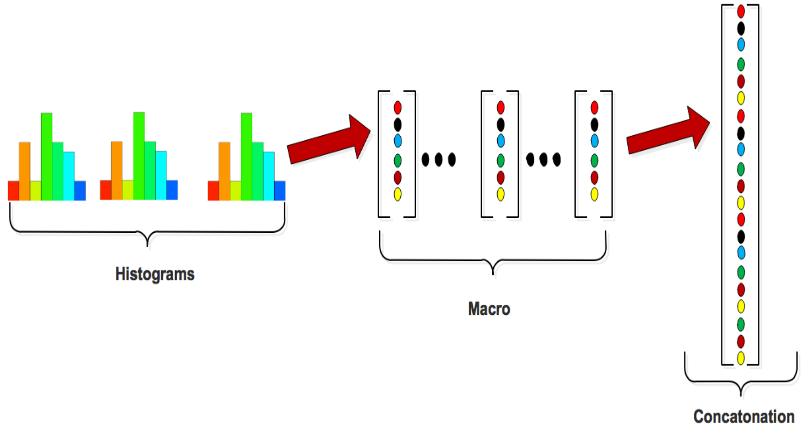


Macro-pattern



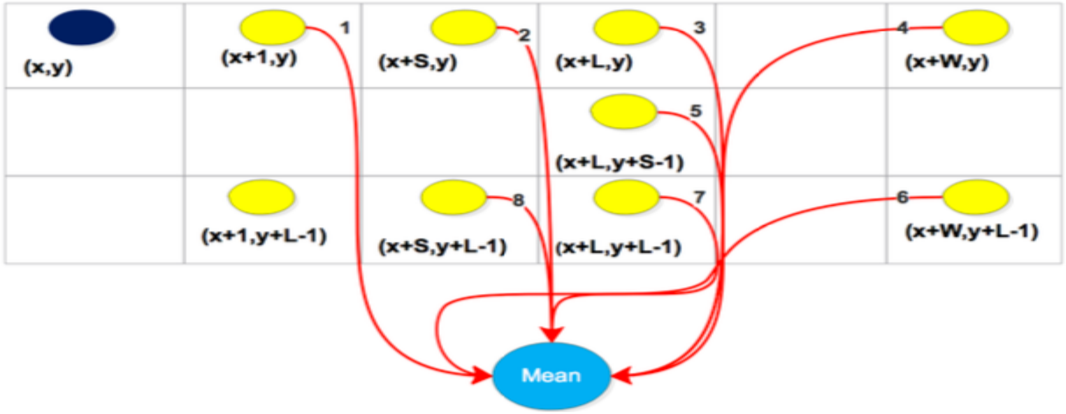


- Better representation.
- Captures better information.



Micro-pattern

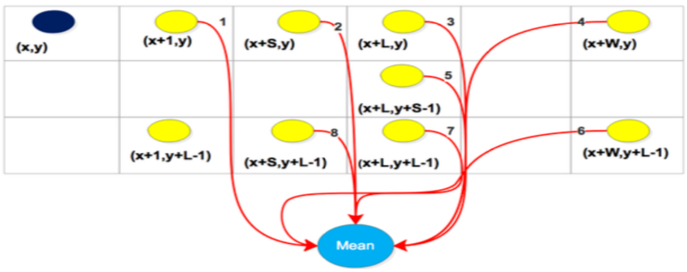
- Apply a noise filter φ

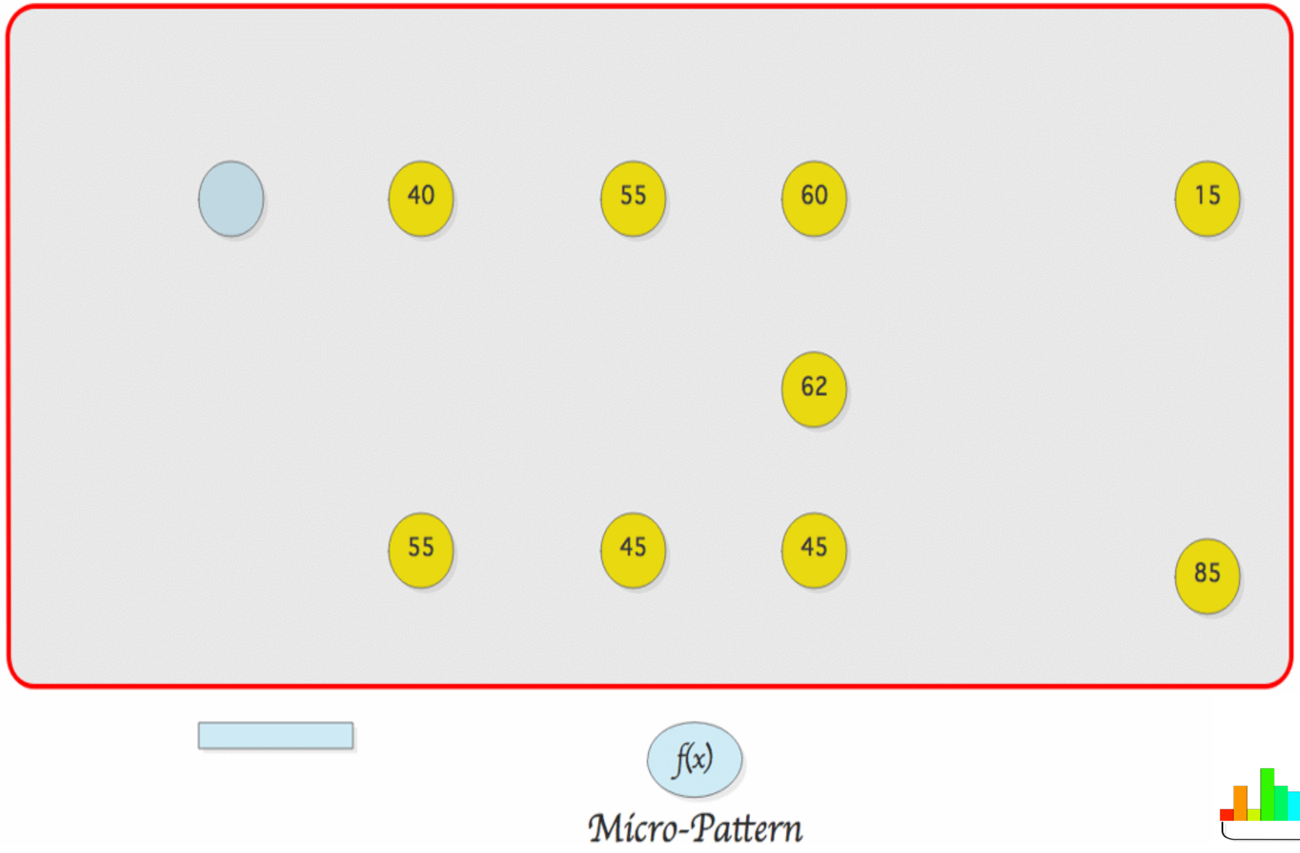


$$f_{dt} = \sum_{n=0}^{7} f_{d,n} - \mu_D$$

$$W_i = \begin{cases} 1, & \text{if } k - \mu_D > 0 \\ 0, & \text{otherwise} \end{cases}$$

Micro-pattern

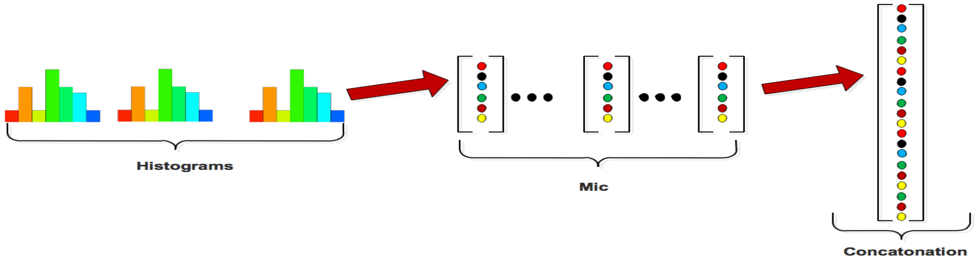




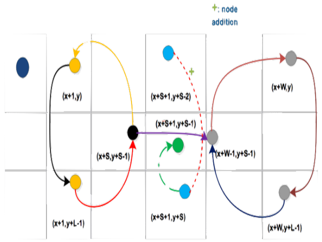
We use the mean instead of median because the values of the nodes are closer to the mean and not the median.

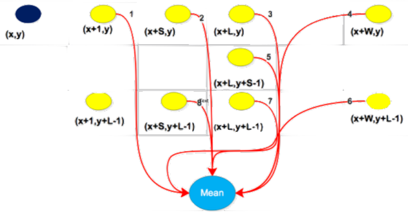
$$f_{dt} = \sum_{n=0}^{7} f_{d,n} - \mu_D$$

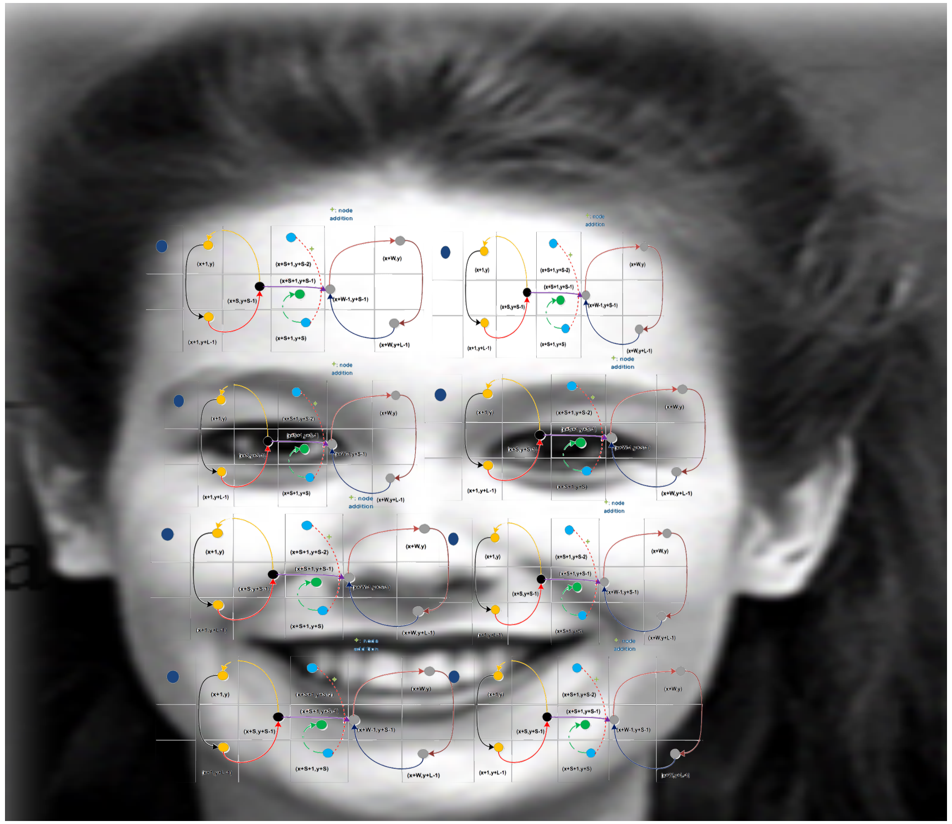
$$W_i = \begin{cases} 1, & \text{if } k - \mu_D > 0 \\ 0, & \text{otherwise} \end{cases}$$

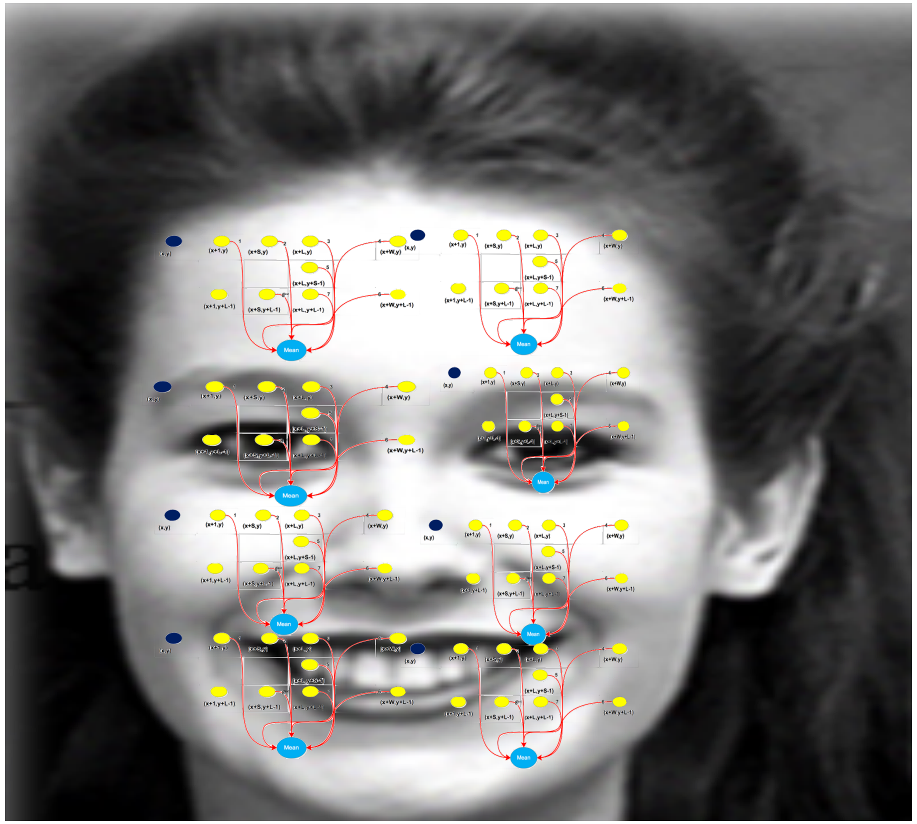


Mic-Macro


+: node addition
(x+1,y)
(x+S+1,y+S-2)
(x+W,y)
(x+S+1,y+S-1)
(x+S,y+S-1)
(x+W-1,y+S-1)
(x+1,y+L-1)
(x+S+1,y+S)
(x+W,y+L-1)


(x,y)
(x+1,y)
(x+S,y)
(x+L,y)
(x+W,y)
(x+L,y+S-1)
(x+1,y+L-1)
(x+S,y+L-1)
(x+L,y+L-1)
(x+W,y+L-1)
Mean



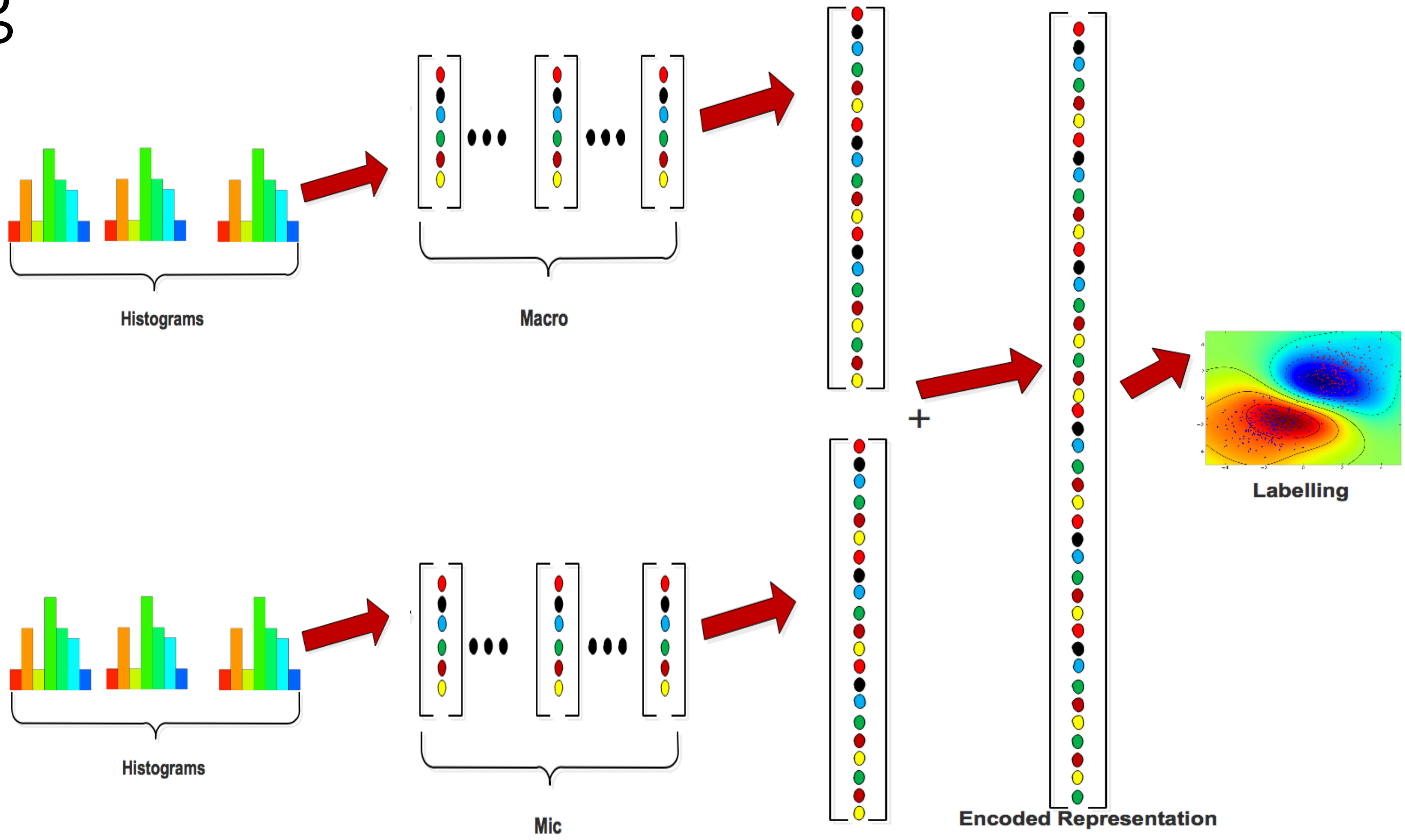


Labelling

Normalize

$$H_t^{'} = \sum_{i=0}^{N} (h_{ti} - \mu_{h_{ti}} / \sigma_{hti}))$$

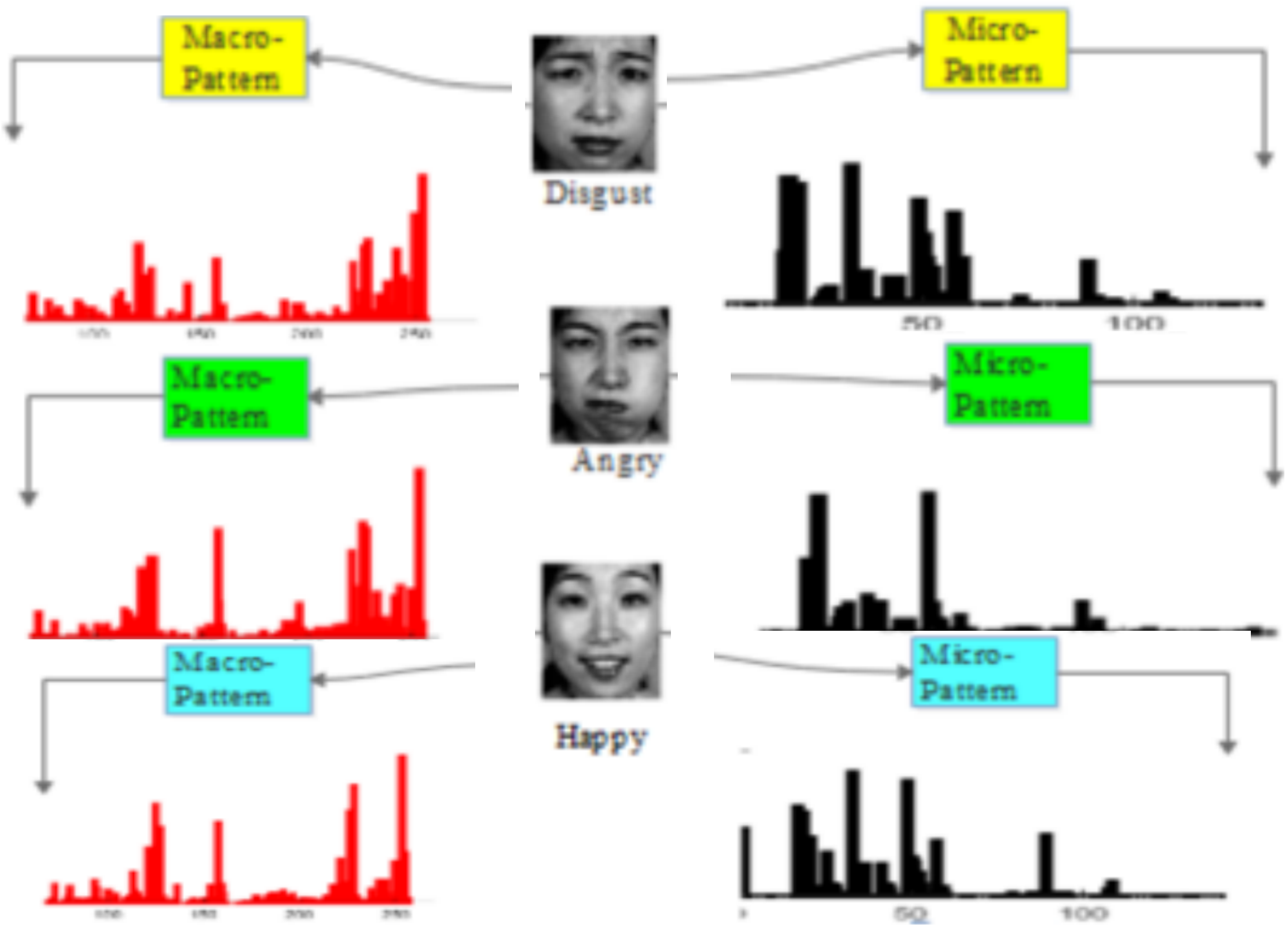
$$H_e^{'} = \sum_{i=0}^{N} (h_{ei} - \mu_{h_{ei}} / \sigma_{hei}))$$



Outline

- Introduction
- Framework
- Mic-Macro patterns
- Experiments results
- Conclusions

Experiments and results



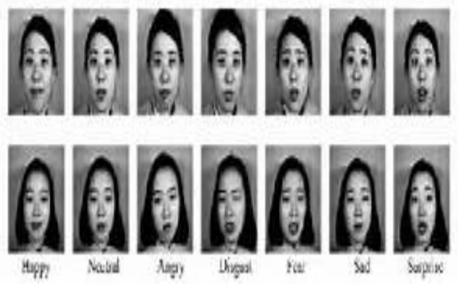
Unique Pattern

Unique Pattern

Unique Pattern

Experiments and results

- Dataset:
  - Japanese Female Facial Expression (JAFFE)
  - Cohn-Kanade (CK)



- Evaluation:
  - k-fold cross validation.
  - Leave one person out validation

experiments: jaffe

Using different radius R

Radius	RBF-kernel (%)	Linear-kernel (%)	Polynomial-kernel (%)
2	90.4762	86.6667	95.2381
3	95.2381	90.4762	95.2381
4	95.2381	97.6190	97.6190
5	**97.6190**	**97.6190**	**97.6190**
6	97.6190	95.2381	95.2381

10-fold cross validation

Ref.	Algorithm	Recognition rate %
[25]	BoostLBP	66.8
[9]	Gabor filter	71.9
[25]	LBP	72.4
[25]	BoostLBP	74.2
[26]	BoostLBP	81.0
[36]	LBP	83.25
[34]	LBP	85.71
[1]	SLGS	88.09
[17]	ASM	89.5
[26]	Gabor filter	91.0
[26]	Gabor filter	91.9
[10]	DCT+Gabor+Wavelets+Gasussian Distribution	93.40
[8]	LBP	93.80
[4]	LPQ+es-LBP-s	96.19
Proposed	micro+macro	**97.61**

Experiments: JAFFE

Using different radius R

Radius	RBF-kernel (%)	Linear-kernel (%)	Polynomial-kernel (%)
2	54.7619	61.9048	55.6324
3	67.4242	70.4969	67.4242
4	72.0779	71.9697	72.0779
5	**79.6066**	**80.9524**	**77.2257**
6	72.8778	72.87781	71.4286

Leave one out cross validation

Ref.	Algorithm	Recognition rate %
[37]	LBP	53.8
[30]	Gabor histogram	58.7
[11]	LBP	62.9
[3]	CT	63.81
[37]	LBP	65.7
[37]	LBP	65.71
[1]	SLGS	66.66
[3]	ASM	68.5
[31]	LBP+Gabor	70.0
[32]	Gabor filter	72.0
[33]	MB-LGBP	74.18
[4]	LPQ+es-LBP-s	76.67
Proposed	micro+macro	**80.95**

Experiments and results

- 10-fold cross-validation on CK:
  - 98%

- Processing Time:
  - We used Matlab R2016b running on windows 10 with Intel Core i7 CPU at 3.60 GHz. The average computation time to extract the micro- and macro-features for one facial image is **0.097 seconds**.

Outline

- Introduction
- Framework
- Mic-Macro patterns
- Experiments results
- Conclusions

Conclusions

- Introduced a framework for facial expression.
- Encoded Mic-Macro patterns from salient patches.
- Proposed method outperforms existing method achieving 80.95% on leave-one-out cross validation.

Thank You!