In The Name of Allah, The Most Beneficent, The Most Merciful

Changes in Facial Expression as Biometric: A Database and Benchmarks of Identification

Rain Eric Haamer, Kaustubh Kulkarni , Nasrin Imanpour, Mohammad A. Haque, Egils Avots, Michelle Breisch, Kamal Nasrollahi, Sergio Escalera, Cagri Ozcinar, Xavier Baro, Ahmad R. Naghsh-Nilchi, Thomas B. Moeslund, Golamreza Anbarjafari

iCV, CVC, UB, UAB, VAP, UI, TCD, OUC, GoSwift, AAU

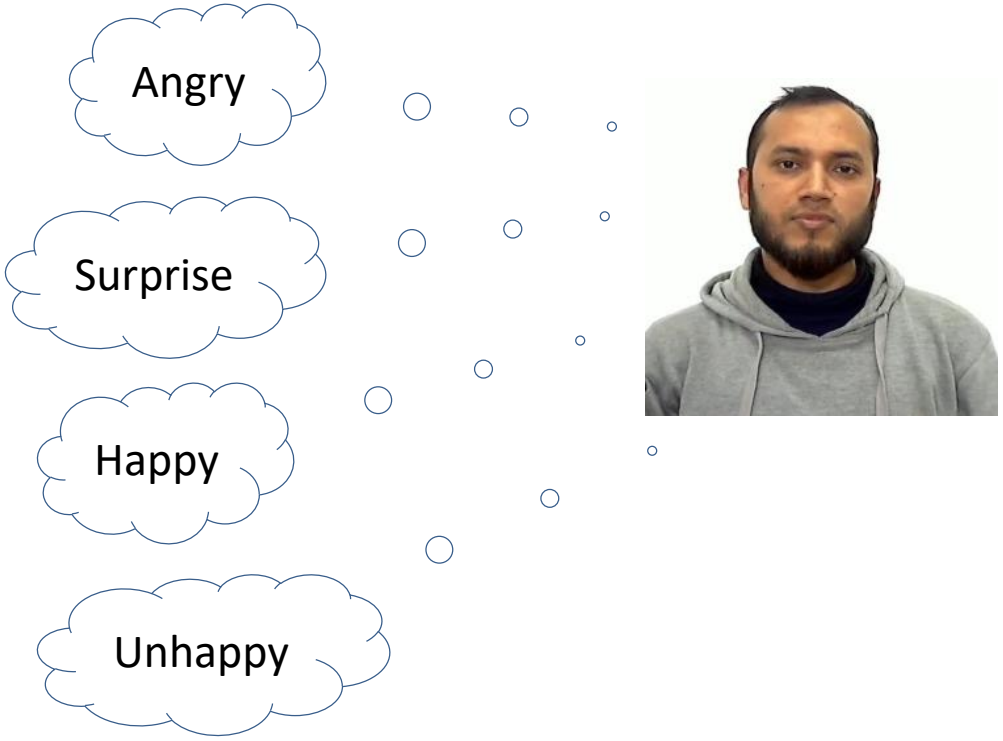
Corresponding author's email: shb@icv.tuit.ut.ee

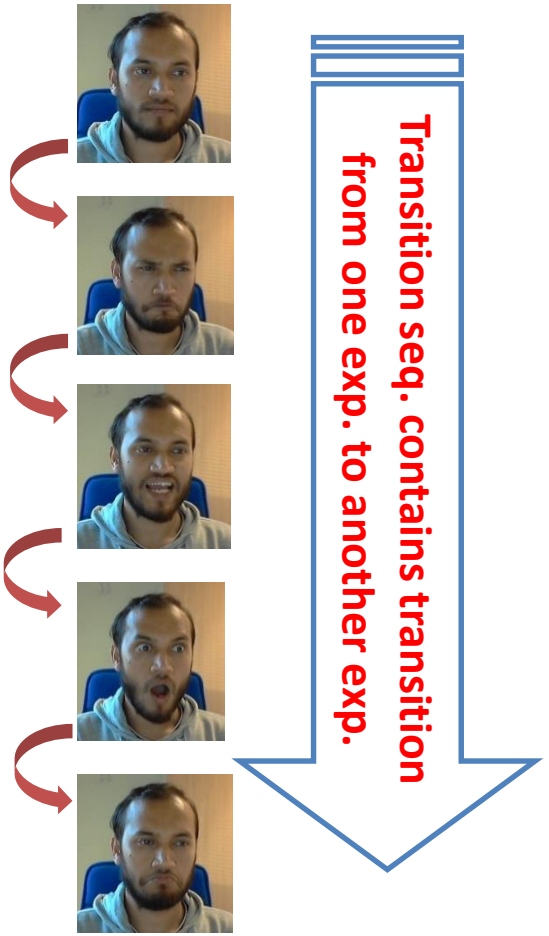
Contents

- Facial expression and transition frames
- Facial expression as biometrics
- A new database
- Exploitation of spatio-temporal information with multimodality
- Remarks to our contributions

Facial Expression and Transition of Expressions

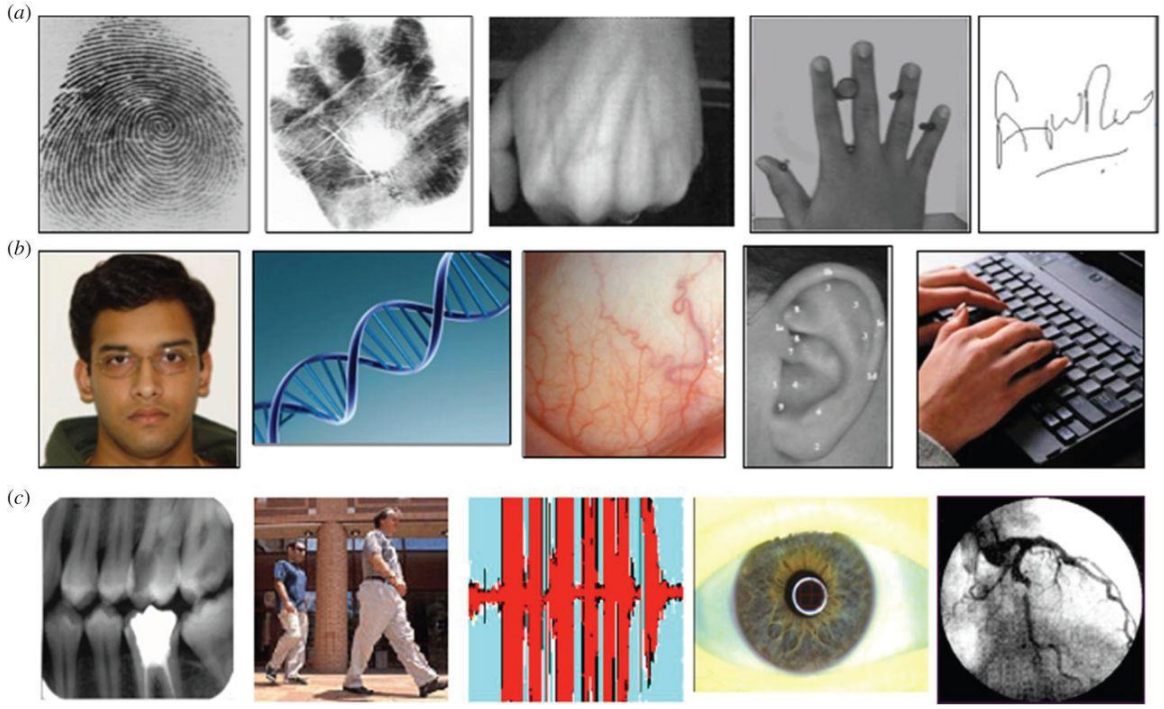
- Human face





Biometrics

- Application-
  - Security systems
  - Forensic analysis
  - Medical record keeping and monitoring
- Traits-
  - Iris image
  - Fingerprint
  - Voice
  - Facial image
  - Hand geometry

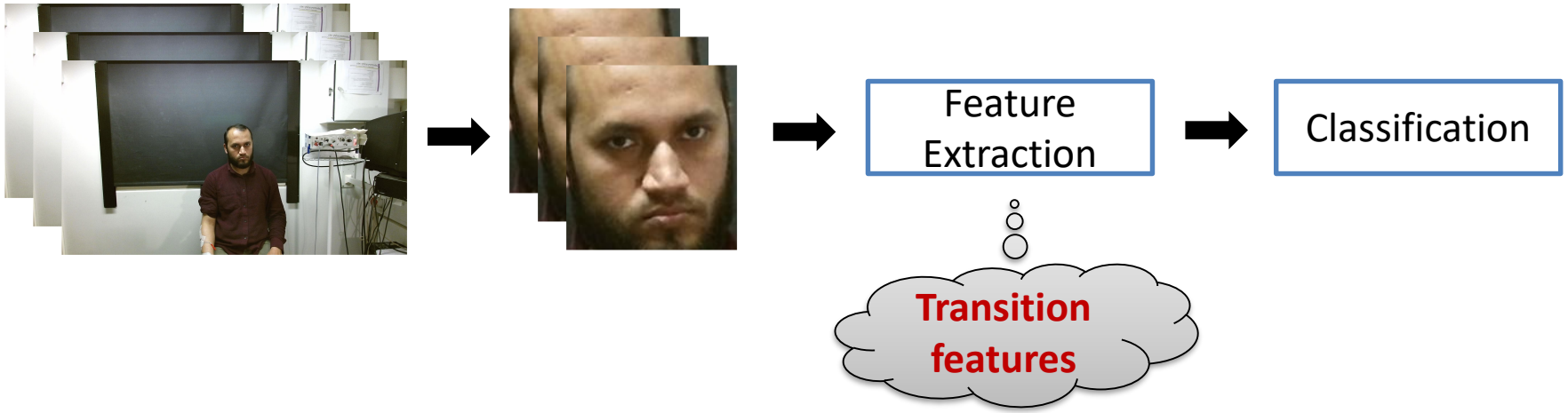


Facial Expression as Biometric

- Studies reported that general facial expressions like sad, anger, happy, etc. can be a biometric or soft-biometric

Ref: Kyong I Chang, Kevin W Bowyer, and Patrick J Flynn. Multiple nose region matching for 3d face recognition under varying facial expression. PAMI, 28(10):1695–1700, 2006.

A New Direction!



- Utilizing facial dynamics that are visible during changes of facial expressions from one motion to the other, to assist regular face recognition systems

Contribution of the paper

- We present
  - Developing the first sequential facial expression transition dataset for face recognition
  - Investigating the use of spatio-temporal emotion features to recognize users
  - Investigating on use of spatio-temporal transition features to recognize users
  - Employing late fusion of emotion and transition systems for boosting recognition rate

The Database

Attributes of the New Expression Transition Database

- **2 sequences for each of the 61 subjects**
- **5 shots for each sequence**

Expression sequences	N-H-U-A-S and N-A-H-S-U

RECORDING SPECIFICATIONS

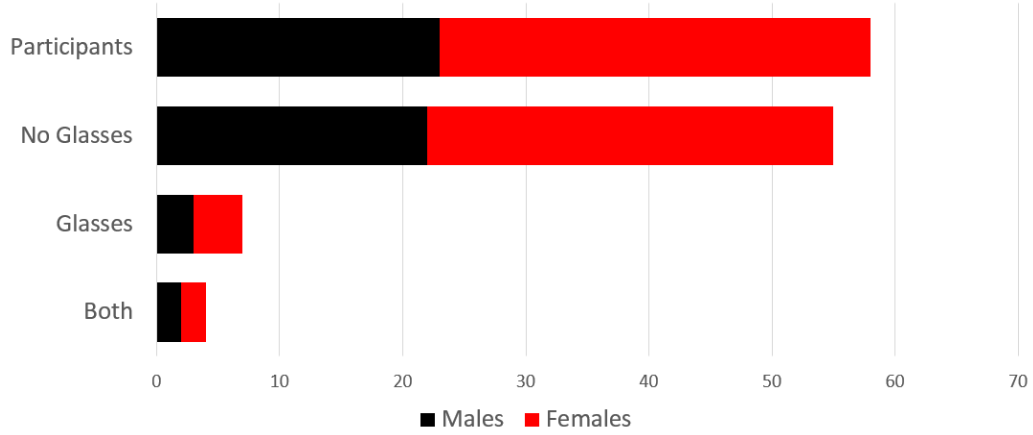
Videos	630
Expressions	5
Mean Duration	15 s
Resolution	1920 x 1080
FPS	25
File type	MTS

PARTICIPANT DEMOGRAPHIC

	Participants	Glasses
Females	39	4
Males	22	1
Total	61	5

PARTICIPANT AGE GROUPS

	Age (y)
Mean	23.64
STD	5.89
Min	18
Max	54



Samples for Two Recorded Seq.

Our Experiments

Methodology

- **5 Frames Corresponding to peak emotion**
- **4 pairs of start and end frames of transition**

- Three experiments:
  - Peak emotion frames
  - Transitoin
  - Fusion of peak emotoin and transition

- Learning:
  - Geomatric features from facial landmarks
  - Deep learning by CNN and LSTM

The Preprocessing of the Database

Raw Frame to Face Alignment

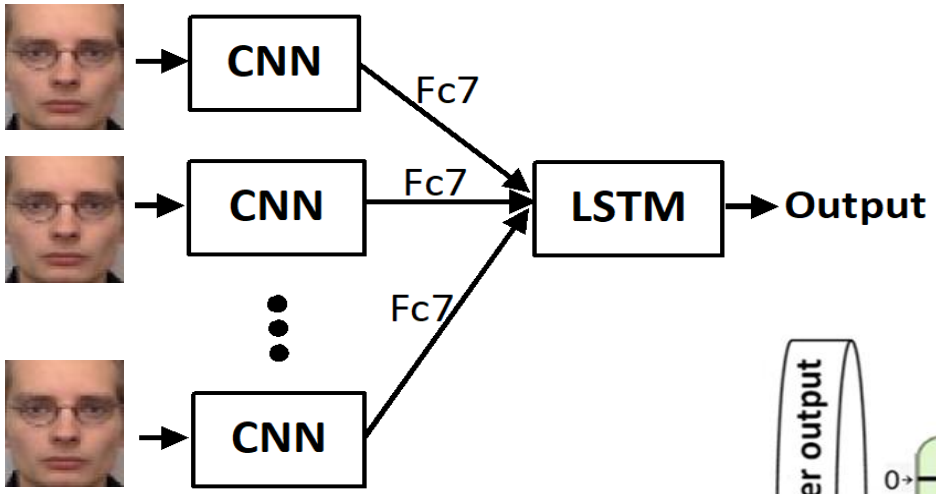
- **5 Frames Corresponding to peak emotion**
- **4 pairs of start and end frames of transition**

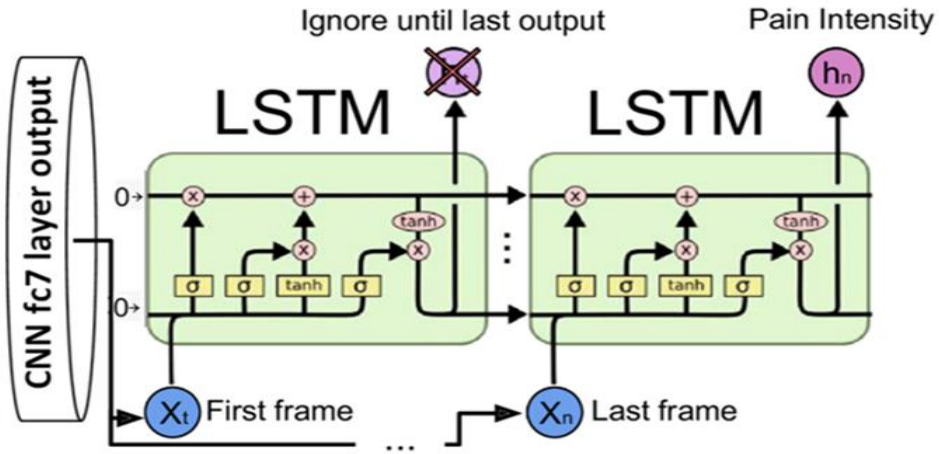
- L. Wolf, T. Hassner, and I. Maoz. Face recognition in unconstrained videos with matched background similarity. In CVPR 2011, pages 529–534, June 2011

- V. Kazemi and J. Sullivan. One millisecond face alignment with an ensemble of regression trees. In CVPR, pages 1867–1874, June 2014.

Exploiting Spatio-Temporal Information

The block diagram of CNN-LSTM based human identification system

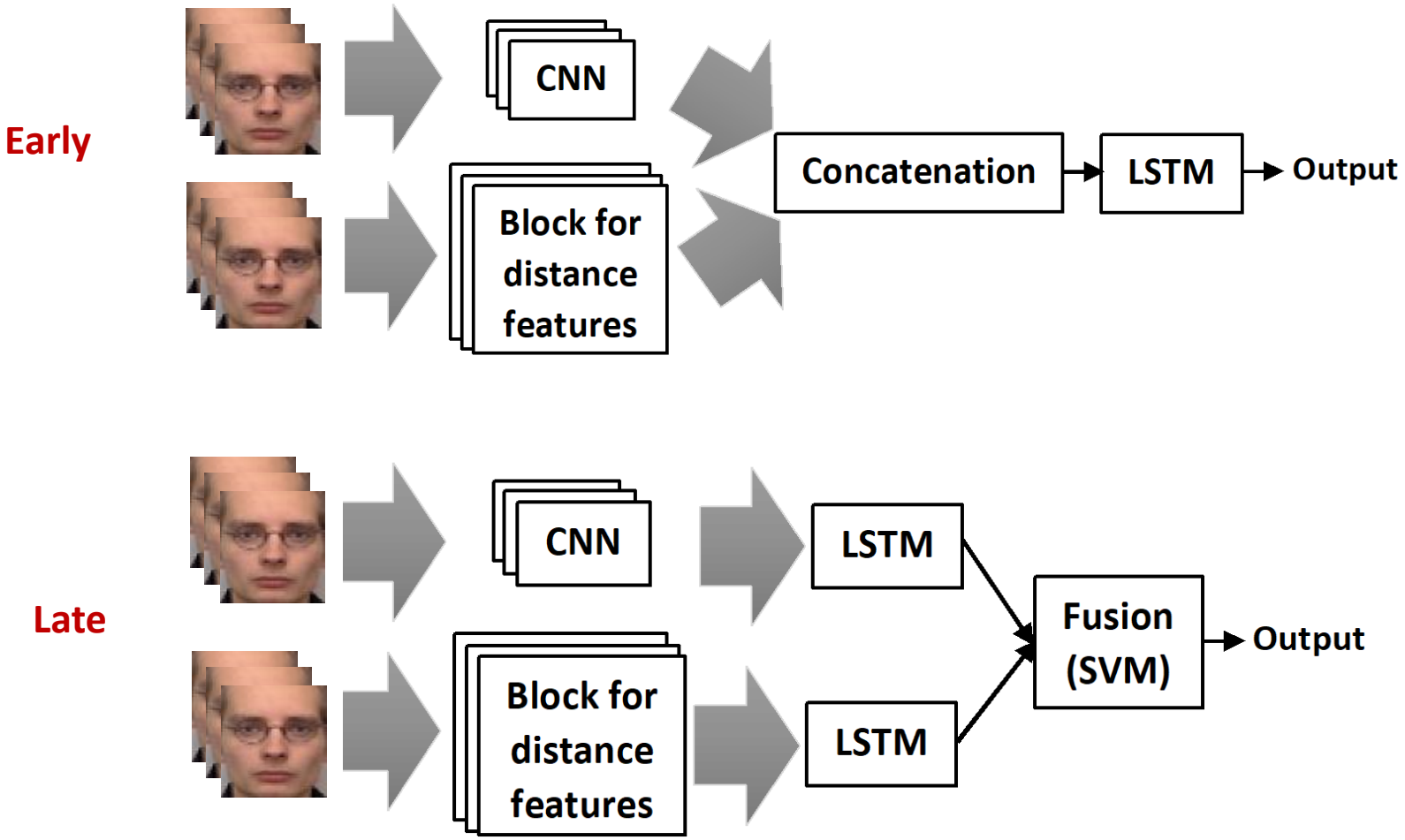




Architecture of the hybrid CNN+LSTM deep learning framework

Exploiting Multimodality

Early and Late Fusions



The Baseline!

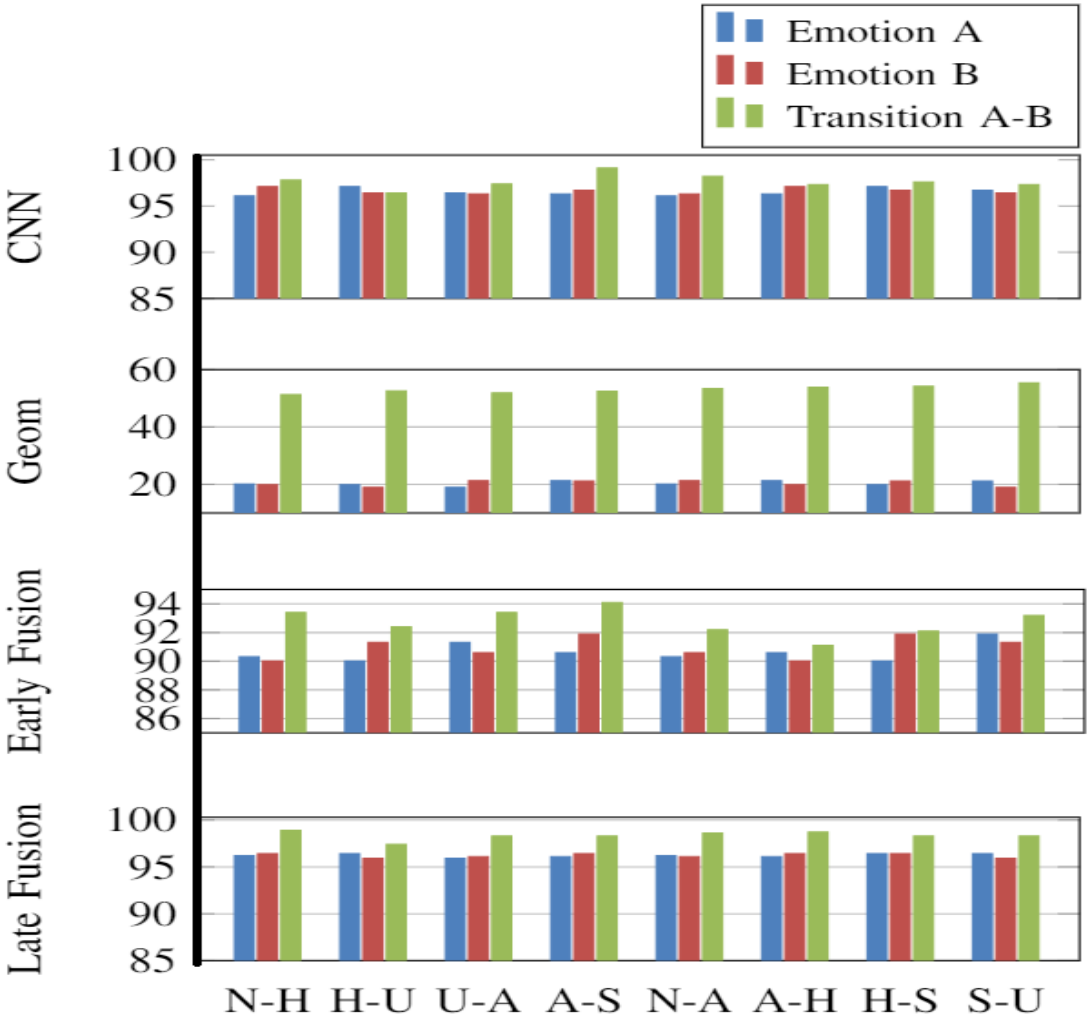
Peak Emotion based Recognition

Method	N	H	U	A	S	Mean
CNN	96.1	97.1	96.4	96.3	96.7	96.5
Geom-SVM	20.1	19.9	18.9	21.2	21.1	20.2
Early Fusion	90.3	90.0	91.3	90.6	91.9	90.8
Late Fusion	96.2	96.4	95.9	96.1	96.4	96.2

Transition-based Recognition

Method	N-H	H-U	U-A	A-S	N-A	A-H	H-S	S-U	Mean
CNN-LSTM	97.8	96.4	97.4	99.1	98.2	97.3	97.6	97.3	97.6
Geom-LSTM	51.3	52.6	51.9	52.5	53.4	53.8	54.2	55.3	53.1
Early Fusion	93.4	92.4	93.4	94.1	92.2	91.1	92.1	93.2	92.7
Late Fusion	98.9	97.4	98.3	98.3	98.6	98.7	98.3	98.3	98.4

A Closer Look to the Transition-based Recognition



Late Fusion of Peak Emotion and Transition based Features

Late fusion Method	N-H	H-U	U-A	A-S	N-A	A-H	H-S	S-U	Mean
P(CNN)+T(CNN+LSTM)	97.6	96.1	97.5	98.7	97.6	97.9	97.2	97.6	97.5
P(CNN)+T(Geom+LSTM)	76.4	77.8	75.1	76.3	78.9	75.8	77.3	77.9	76.9

Conclusions

- When doing late fusion of the geometric and the CNN pipelines, we observe that this pipeline gives the best accuracy of all experiments. This means, there is some complimentary information to the appearance features in the geometric features that helps recognition.

- The final experimental result showed that the transition frames outperform the peak emotion frames in face recognition.

- As a drawback, one must notice that the subjects acts the emotion with intensity.

MISC

- The database is available here:
  - Corresponding author's email: shb@icv.tuit.ut.ee

- An implementation of CNN+LSTM available here:
  - https://github.com/prlz77/LSTM-on-CNN


Q
&
A