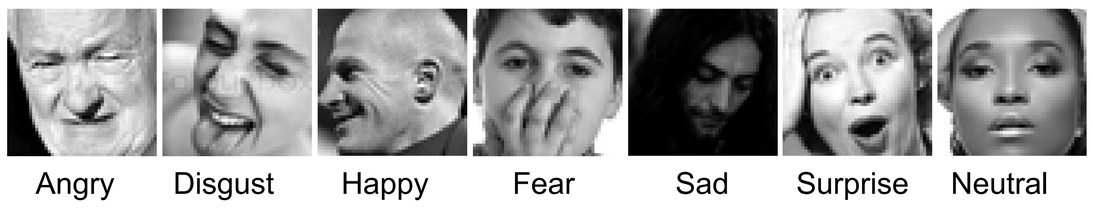
Automated Recognition of Emotions Based on Images of Facial Expressions

Robert Berszan, Krisztian Buza
Biointelligence Group, Department of Mathematics-Informatics
Sapientia Hungarian University of Transylvania, Targu Mures, Romania
robertberszan@yahoo.com, buza@biointelligence.hu

Abstract

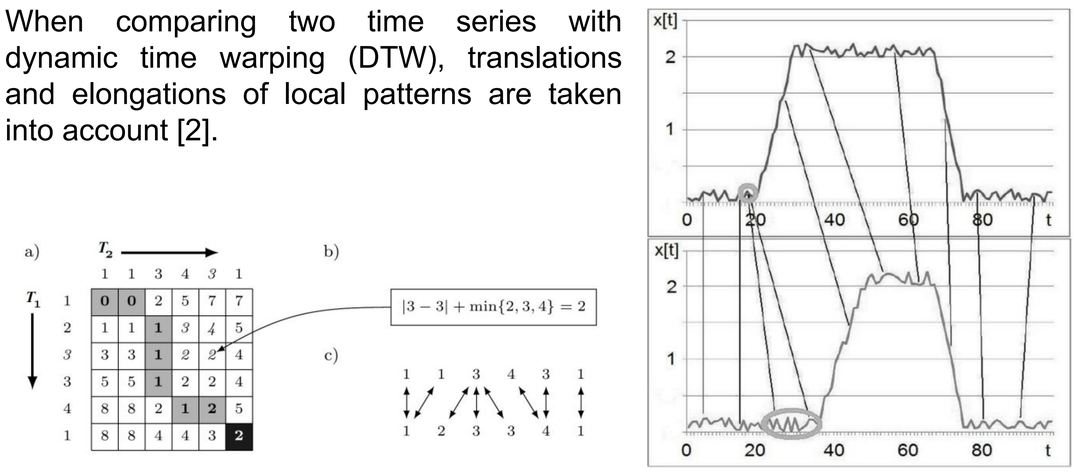
Personal computers are involved in various cognitive science studies, e.g. to measure human subjects' preferences or response times quantitatively. While the subjects' emotions may be essential to explain the observations, objective assessment of emotions in such studies is inherently difficult. However, state-of-the-art computers are equipped with cameras that could be used to capture an image of the subject's face based on which her/his emotions may be recognized automatically which may provide objective assessment of emotions in large-scale studies. In order to take the first steps into this direction, we considered a publicly available image dataset [1] and compared distance-based machine learning techniques for the automated recognition of the emotions. The dataset contains 35882 grayscale images (resolution: 48x48). Each image shows the face of a human subject and it is annotated with one of the following emotions: angry, disgust, fear, happy, sad, surprise and neutral. Besides considering well-known distance measures, such as Euclidean distance, Manhattan distance or cosine distance, we used dynamic time warping (DTW) in a genuine way to compare the images of facial expressions and to recognize emotions.

Data

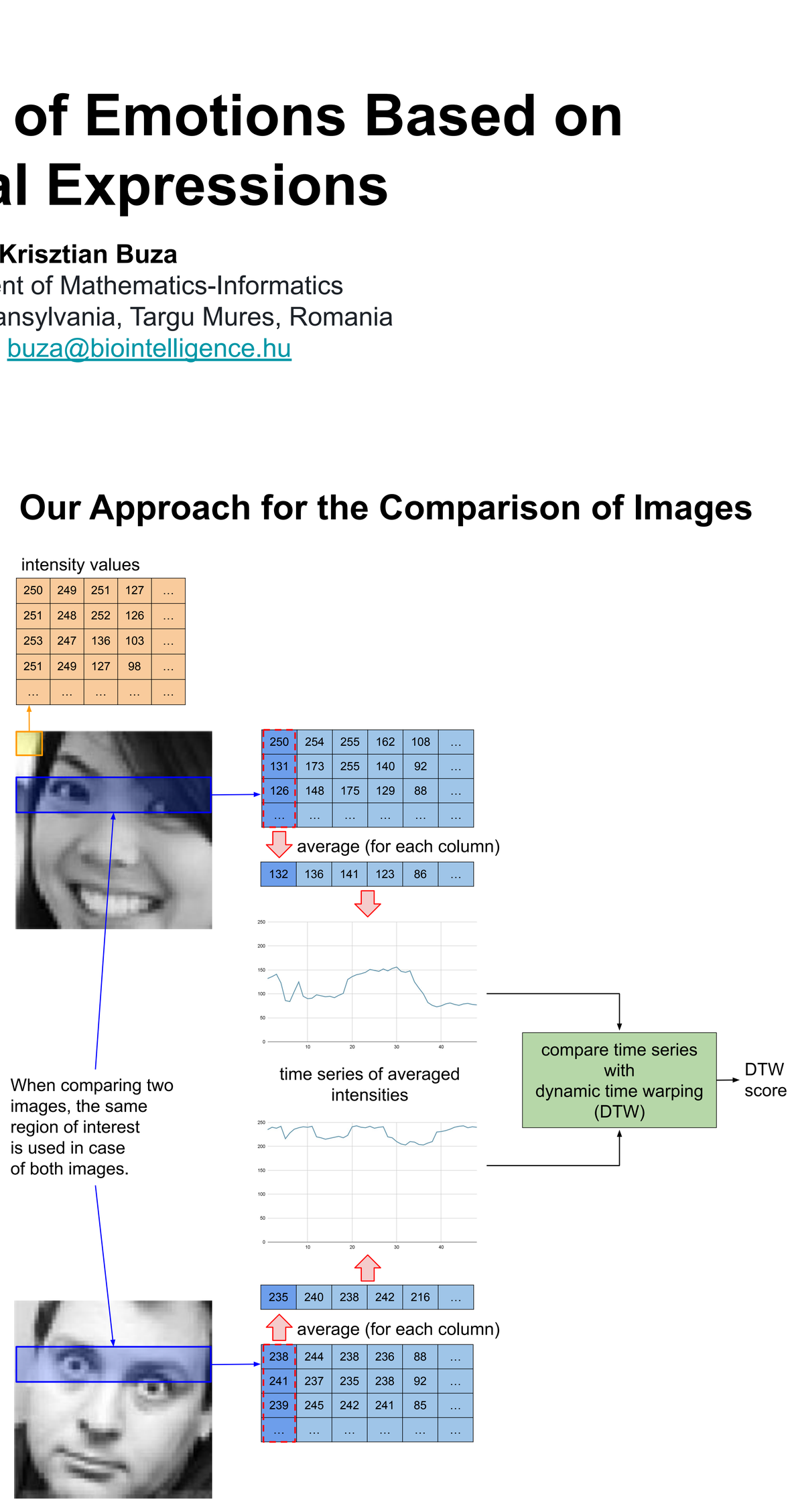


Background: Dynamic Time Warping (DTW)

When comparing two time series with dynamic time warping (DTW), translations and elongations of local patterns are taken into account [2].



Our Approach for the Comparison of Images



We defined six regions of interest, thus we calculated six DTW scores for a pair of images. We use the average of these scores to quantify the similarity between the two images.

For regions of interest with vertical orientation, we calculate the average for each row, instead of calculating it for each column.

Results

nearest neighbor classification, 10-fold cross-validation

No.	Distance	Accuracy (%)
1	Euclidean	59.99 ± 0.74
2	Manhattan	61.03 ± 0.66
3	Cosine	59.74 ± 1.10
4	**Our approach**	**62.95 ± 0.54**

References

[1] Facial Expression Data
https://www.kaggle.com/debanga/facial-expression-recognition-challenge

[2] K. Buza (2018): Time Series Classification and its Applications
8th International Conference on Web Intelligence, Mining and Semantics