# Automatic Estimation of Self-Reported Pain by Interpretable Representations of Motion Dynamics

Benjamin Szczapa<sup>1,3</sup>, Mohamed Daoudi<sup>2</sup>, Stefano Berretti<sup>3</sup>, Pietro Pala<sup>3</sup>, Alberto Del Bimbo<sup>3</sup> and Zakia Hammal<sup>4</sup>

<sup>1</sup>Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France <sup>2</sup>IMT Lille Douai, Univ. Lille, CNRS, UMR 9189 CRIStAL, F-59000 Lille, France <sup>3</sup>Department of Information Engineering, University of Florence, Italy <sup>4</sup>Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA



- 1. Context for Pain Estimation
- 2. Presentation of our Approach
- 3. Experimental Results
- 4. Conclusion

## Context for Pain Estimation

• Pain is an unpleasant sensory

- Pain is an unpleasant sensory
- Pain is an emotional experience

- Pain is an unpleasant sensory
- Pain is an emotional experience
- Pain is **subjective**

- Pain is an unpleasant sensory
- Pain is an emotional experience
- Pain is subjective
- Difficult to obtain

- Pain is an unpleasant sensory
- Pain is an emotional experience
- Pain is subjective
- Difficult to obtain
  - Children

- Pain is an unpleasant sensory
- Pain is an emotional experience
- Pain is **subjective**
- Difficult to obtain
  - Children
  - Disabled patients

- Pain is an unpleasant sensory
- Pain is an emotional experience
- Pain is **subjective**
- Difficult to obtain
  - Children
  - Disabled patients
  - Patients requiring breathing assistance

• Most methods estimate pain at frame-level (PSPI, based on FACS)

- Most methods estimate pain at frame-level (PSPI, based on FACS)
- Most methods use deep learning approaches (RNN)

- Most methods estimate pain at frame-level (PSPI, based on FACS)
- Most methods use deep learning approaches (RNN)
- Few work on sequence-level pain estimation (VAS)

## **Proposed Method**



Figure 1: Overview of our method for action recognition

- $\cdot$  Geometry based approach to estimate self-reported pain
- Facial dynamics based on Gram matrix computation and trajectory modeling on the Riemannian manifold of positive semi-definite (PSD) matrices of fixed rank
- The manifold  $S^+(d, n)$  of PSD matrices is endowed with an optimized metric for 2D
- A recent curve fitting method is used to smooth trajectories on the manifold
- The use of Global Alignment Kernel for temporal alignment, instead of DTW

## Presentation of our Approach

• Each sequence is characterized by a set of landmarks  $\{Z_0, \ldots, Z_{\tau}\}$ , where  $\tau$  is the number of frames of the sequence

- Each sequence is characterized by a set of landmarks  $\{Z_0, \ldots, Z_{\tau}\}$ , where  $\tau$  is the number of frames of the sequence
- Each configuration matrix  $Z_i$   $(1 \le i \le \tau) \in \mathbb{R}^{n \times d}$  encodes the position of *n* landmarks in *d* dimensions

- Each sequence is characterized by a set of landmarks  $\{Z_0, \ldots, Z_{\tau}\}$ , where  $\tau$  is the number of frames of the sequence
- Each configuration matrix  $Z_i$   $(1 \le i \le \tau) \in \mathbb{R}^{n \times d}$  encodes the position of *n* landmarks in *d* dimensions
- Velocities are computed as the magnitude of the displacement between two consecutive landmark configurations Z<sub>i</sub> and Z<sub>i+1</sub>

- Each sequence is characterized by a set of landmarks  $\{Z_0, \ldots, Z_{\tau}\}$ , where  $\tau$  is the number of frames of the sequence
- Each configuration matrix  $Z_i$   $(1 \le i \le \tau) \in \mathbb{R}^{n \times d}$  encodes the position of *n* landmarks in *d* dimensions
- Velocities are computed as the magnitude of the displacement between two consecutive landmark configurations *Z<sub>i</sub>* and *Z<sub>i+1</sub>*
- The final facial representation A is the concatenation of the landmark coordinates and velocities (size 2n)

Gram matrix representation as the inner product of each facial configuration matrix:

$$G = AA^{T} = \left\langle p_{i}, p_{j} \right\rangle, \quad 1 \le i, j \le 2n .$$
(1)

We consider the Riemannian geometry of the space  $S^+(d, 2n)$  of  $2n \times 2n$  PSD matrices of rank d. We define a Riemannian metric, resulting in the following distance between PSD matrices:

$$d(G_i, G_j) = tr(G_i) + tr(G_j) - 2tr\left(\left(G_i^{\frac{1}{2}}G_jG_j^{\frac{1}{2}}\right)^{\frac{1}{2}}\right)$$
(2)

### **Riemannian Geometry**

This distance can be expressed in terms of the facial configurations  $A_i, A_j \in \mathbb{R}^{m \times d}_*$  as follows:

$$d(G_i, G_j) = \min_{Q \in \mathcal{O}_d} ||A_j Q - A_i||_F$$
(3)

The optimal solution is  $Q^* := VU^T$ , where  $A_i^T A_j = U \Sigma V^T$  is a singular value decomposition.

In the specific case of 2D landmarks, the distance can be reformulated as:

$$d(G_i, G_j) = \operatorname{tr}(G_i) + \operatorname{tr}(G_j) - 2\sqrt{(a+d)^2 + (c-b)^2}$$
(4)  
where  $A_i^T A_j = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ .

## Sequence Alignment and Pain Estimation

Global Alignment Kernel (GAK) is used instead of Dynamic Time Warping for sequence alignment.

- We compute the matrix *D* that contains the distances between all the elements of two sequences
- We compute a kernel *k* using the halved Gaussian Kernel on this distance matrix *D*
- We compute the similarity score between two sequences

we define a zeros matrix *M* of size  $(\tau_1 + 1) \times (\tau_2 + 1)$  with  $M_{0,0} = 1$  that will contain the path to the similarity between our two sequences. The terms of *M* are computed as:

$$M_{i,j} = (M_{i,j-1} + M_{i-1,j-1} + M_{i-1,j}) * k(i,j).$$
(5)

The similarity score is the value at  $M_{(\tau_1+1),(\tau_2+1)}$ 

Finally, we build a new matrix *K* of size  $n_{seq} \times n_{seq}$ , where  $n_{seq}$  is the number of sequences in the dataset that contains the similarity scores between all the sequences and is used directly with SVR for self-reported pain estimation.

# Experimental Results

## **UNBC-McMaster Shoulder Pain Archive**

The UNBC-McMaster Shoulder Pain Archive [2] contains:

- 200 facial videos
- 25 subjects
- Both sequence-level (VAS) and frame-level (PSPI) label



**Figure 2:** Example images from the UNBC-McMaster Shoulder Pain Archive and their corresponding landmark coordinates and velocities.

VAS Score	Number of Sequences
0	35
1	42
2	24
3	20
4	21
5	11
6	11
7	6
8	18
9	10
10	2

Table 1: Distribution of the VAS painscores in the UNBC-McMasterShoulder Pain Archive



**Figure 3:** Number of sequences per subjects in the UNBC-McMaster Shoulder pain archive dataset.

- Leave-One-Sequence-Out
- Leave-One-Subject-Out cross validation (subject-independent)
- 5-folds cross validation (subject-independent)

Protocol	% of frames	MAE	RMSE
Logvo-Opo-Soguopeo-Out	25%	2.3166	3.1459
Leave-One-Sequence-Out	100%	2.5291	3.3263
Logvo-Opo-Subject-Out cross validation	25%	2.523	3.2692
Leave-One-Subject-Out closs valuation	100%	2.9176	3.5133
5-fold cross validation	25%	2.4365	3.147
	100%	2.7944	3.5088

Table 2: Results of our method with the three different protocols.

Method	Protocol	Labels for training	MAE
DeepFaceLift [1]	5-fold cross validation	VAS	2.30
RNN-HCRF [3]	random split	VAS & PSPI	2.46
Ours	5-fold cross validation	VAS	2.4365

Table 3: Comparison of our method with state-of-the-art results

[1] DeepFaceLIFT: Interpretable Personalized Models for Automatic Estimation of Self-Reported Pain , [3] Personalized Automatic Estimation of Self-Reported Pain Intensity from Facial Expressions

Conclusion

- We proposed a geometry based approach to estimate self-reported pain
- Method based on facial landmarks for anonymity
- Results can be easily interpreted
- Competitive with state-of-the-art results

Contact: benjamin.szczapa.etu@univ-lille.fr

#### References i

D. Liu, F. Peng, O. O. Rudovic, and R. W. Picard.
 Deepfacelift: Interpretable personalized models for automatic estimation of self-reported pain.

In AffComp@IJCAI, volume 66 of Proceedings of Machine Learning Research, pages 1–16. PMLR, 2017.

P. Lucey, J. F. Cohn, K. M. Prkachin, P. E. Solomon, and I. A. Matthews.

Painful data: The unbc-mcmaster shoulder pain expression archive database.

In Ninth IEEE International Conference on Automatic Face and Gesture Recognition (FG 2011), Santa Barbara, CA, USA, 21-25 March 2011, pages 57–64, 2011.

#### D. L. Martinez, O. Rudovic, and R. W. Picard. Personalized automatic estimation of self-reported pain intensity from facial expressions.

In IEEE Conf. on Computer Vision and Pattern Recognition Workshops CVPR, pages 2318–2327, 2017.