

Chun-Hao Huang<sup>1</sup>, Edmond Boyer<sup>2</sup> and Slobodan Ilic<sup>1</sup>

<sup>1</sup> Computer Aided Medical Procedures (CAMP), TU Munich, Germany

<sup>2</sup> Morpheo, INRIA Grenoble Rhône-Alpes

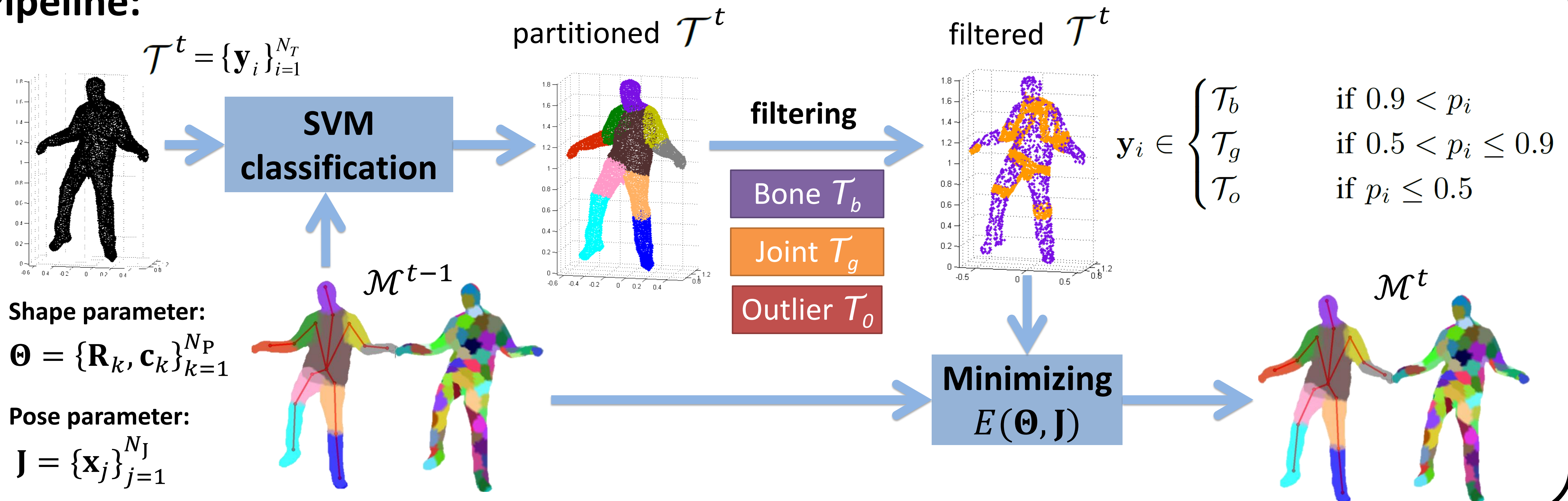
## Motivation:

- Surface-based approaches [4] handle non-rigid deformation better, but do not provide poses.
- Skeleton-based methods [1, 2] significantly reduce the dimension of search space, but usually require 2<sup>nd</sup> stage shape refinement.

## Contribution:

- A SVM-based classification approach to help correspondence search.
- The combination of patch-based surface deformation [4] and bone differential coordinate [3] which recover the pose and the shape at the same time.

## Pipeline:

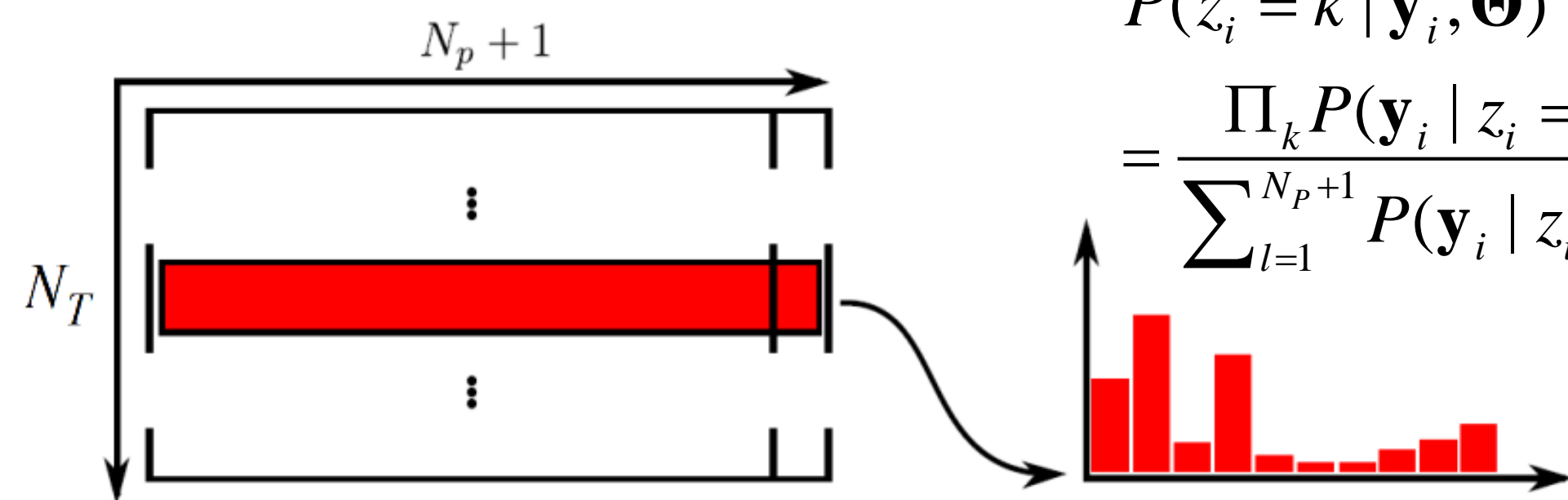


## Energy terms: $E(\Theta, \mathbf{J}) = E_{\text{data}}(\Theta) + E_{\text{rigid}}(\Theta) + E_{\text{bone}}(\Theta, \mathbf{J})$

- $E_{\text{data}}(\Theta)$ : EM algorithm

M-step: minimizing sum of negative log likelihood

E-step:

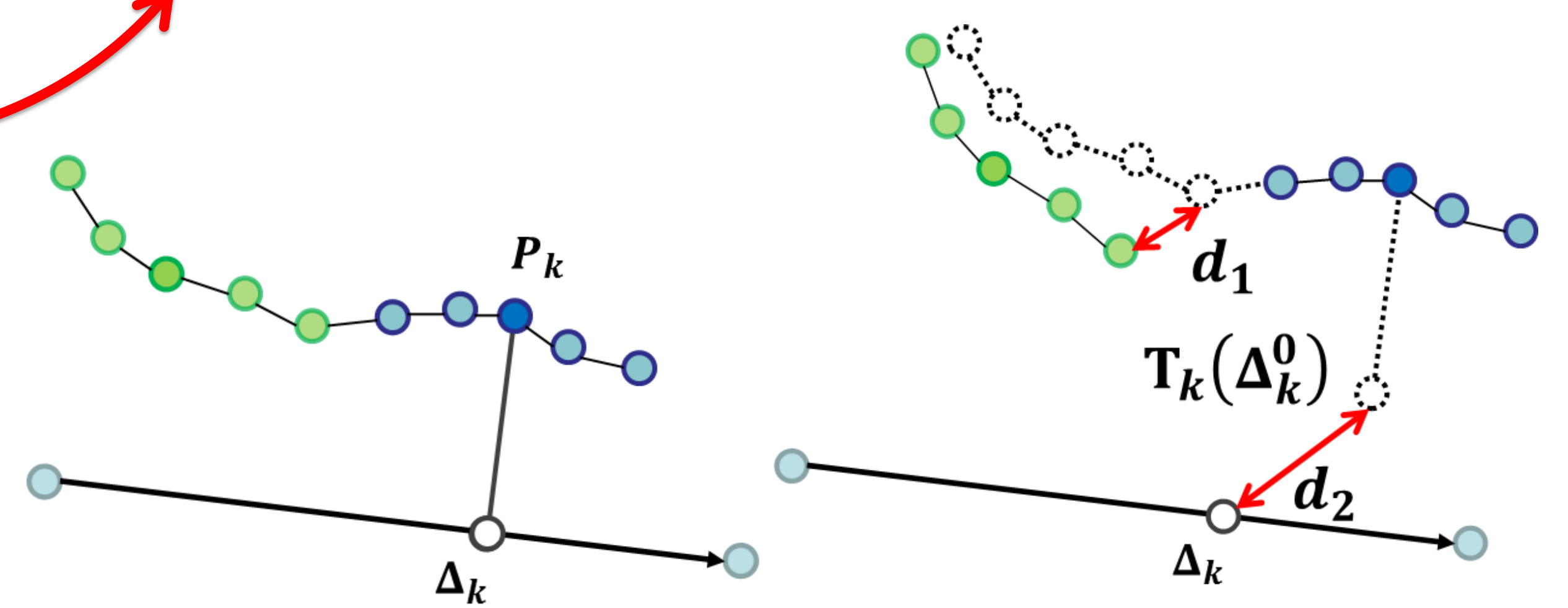


$$E_{\text{data}}(\Theta) = \sum_{i=1}^{N_T} \sum_{k=1}^{N_P+1} w_i^k \left\| \mathbf{y}_i - \mathbf{x}_{v_i^k} \right\|^2$$

where  $w_i^k = \frac{P(z_i = k | \mathbf{y}_i, \Theta)}{\sum_{l=1}^{N_P+1} P(z_i = l | \mathbf{y}_i, \Theta)}$  and  $P(\mathbf{y}_i | \Theta) = \sum_{k=1}^{N_P+1} \Pi_k P(\mathbf{y}_i | z_i = k, \Theta)$ .

- $E_{\text{rigid}}(\Theta)$  &  $E_{\text{bone}}(\Theta, \mathbf{J})$ : regularization terms

$E_{\text{rigid}}(\Theta)$  minimizes  $d_1$  and  $E_{\text{bone}}(\Theta, \mathbf{J})$  minimizes  $d_2$

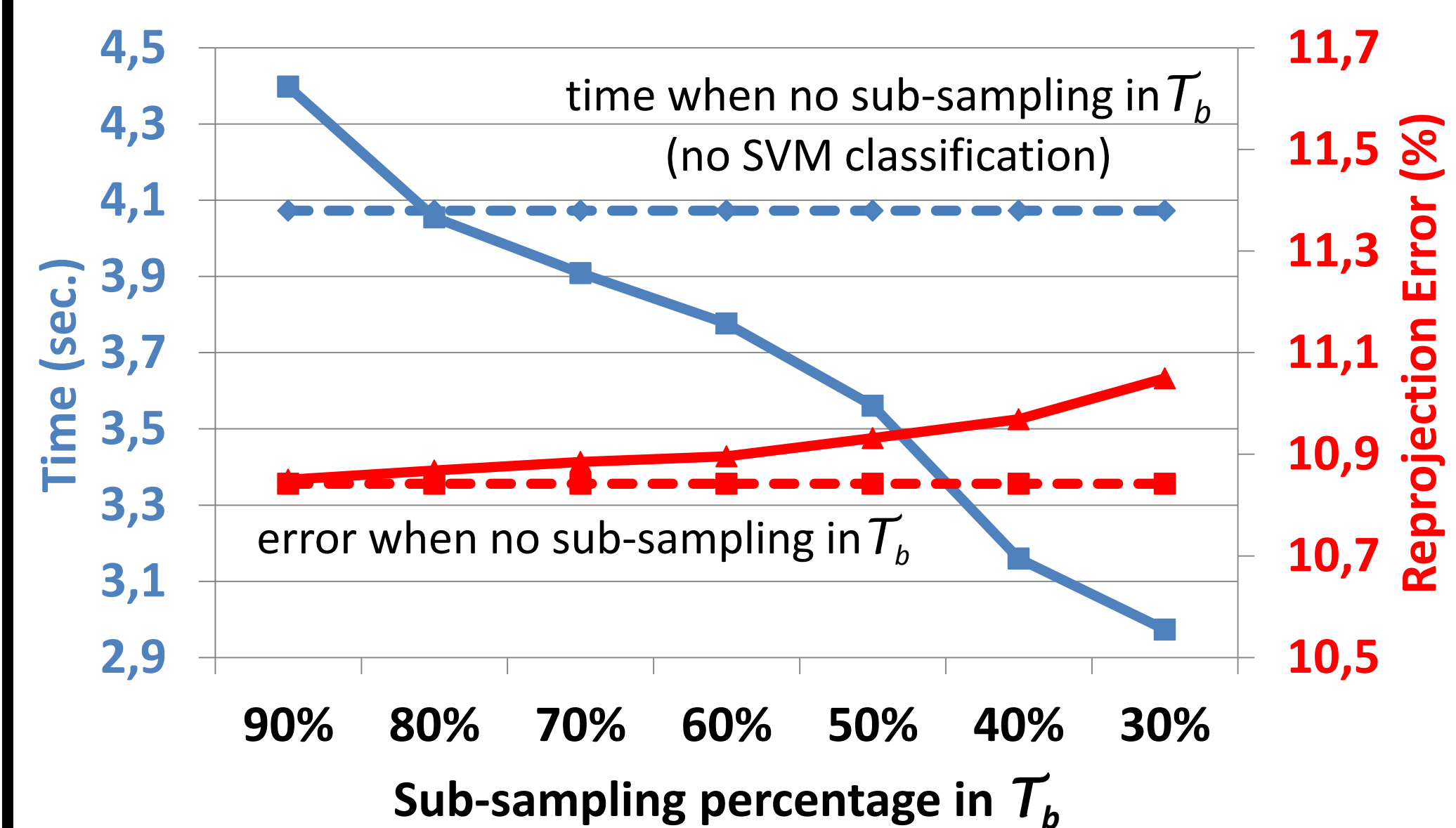


## Experiment Results:

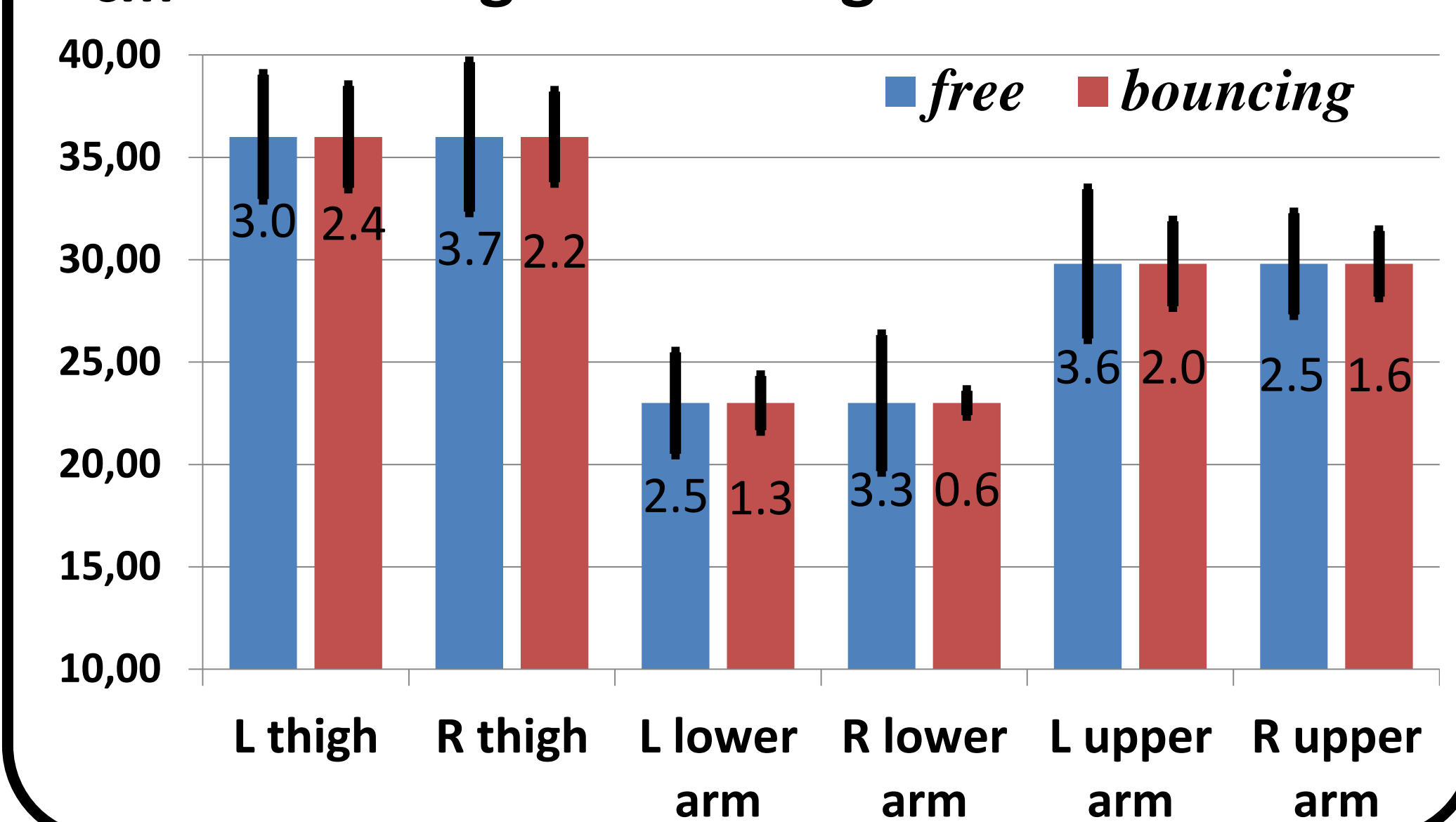
- **Pose:** 70.86mm in *Walking* sequence from HumanEvaII dataset (error < 80mm corresponds to a correct pose [Sigal et al. IJCV'12])
- **Shape:** reprojection error (%)

Sequence	our	Cagniard et al. [4]	inverse kinematic	other approaches
Handstand 1 [1]	15.53	20.13	23.04	7.66 [1]
Wheel [1]	10.28	10.30	14.35	6.42 [1]
Skirt [1]	11.94	12.55	21.43	9.56 [1]
Dance [1]	9.95	9.90	15.01	12.08 [1]
Crane [2]	10.79	11.20	16.33	5.29 [2]
Handstand 2 [2]	12.84	13.97	15.16	5.73 [2]
Bouncing [2]	9.87	9.95	14.64	5.34 [2]
Free [3]	14.12	14.69	-	-

## Error vs. time in Crane sequence [1]



## Average bone length & variation



## Selected References:

- [1] D. Vlastic, I. Baran, W. Matusik, and J. Popovic. Articulated mesh animation from multi-view silhouettes, *TOG*, 2008
- [2] J. Gall, C. Stoll, E. De Aguiar, C. Theobalt, B. Rosenhahn, and H.-P. Seidel. Motion capture using joint skeleton tracking and surface estimation, *CVPR*, 2009
- [3] M. Straka, S. Hauswiesner, M. R. uther, and H. Bischof. Simultaneous shape and pose adaption of articulated models using linear optimization, *ECCV*, 2012
- [4] C. Cagniard, E. Boyer, and S. Ilic. Probabilistic deformable surface tracking from multiple videos, *ECCV*, 2010
- [5] J. Starck and A. Hilton. Surface capture for performance based animation, *Computer Graphics and Applications*, 2007