Supplementary Material for
Compact Matrix Factorization with Dependent Subspaces

Viktor Larsson[1]

Carl Olsson[1,2]

[1]Centre for Mathematical Sciences
Lund University

[2]Department of Signals and Systems
Chalmers University of Technology

{viktorl,calle}@maths.lth.se

1. Spatial Smoothness for Labeling

In this section we consider the incorporation of a pairwise Potts penalty [1] in our energy based model fitting framework. This regularization term, typically referred to as the smoothness term, adds spatial context to the formulation and can resolve ambiguous cluster assignments. This is particularly effective in areas where several good subspaces are available, e.g. in the vicinity of transitions between models.

We use the formulation

$$E(l) = \sum_p \left(\sum_{q \in \mathcal{N}(p)} \mu S_{pq}(l_p, l_q) + D_p(l_p) \right) + \sum_k h_k \delta_k(l). \tag{1}$$

Recall that the data terms $D_p(l_p)$ and the label costs $h_k \delta_k(l)$ jointly encode a trade-off between data fit and model complexity. The additional term $S_{pq}(l_p, l_q)$ penalizes cluster assignments where neighboring trajectories are assigned different clusters. The function S_{pq} assumes the values 0 or 1 and the parameter μ controls the penalty strength.

For the neighborhood system we use the k nearest neighbors. We define the distance between column k and column ℓ as the maximum distance between the elements in their overlapping rows, i.e.

$$\text{dist}\,(M_k, M_\ell) = \max_{W_{ik}=W_{i\ell}=1} |M_{ik} - M_{i\ell}|. \tag{2}$$

If the two columns do not have any overlapping observations we define the distance to be infinite. In the following experiments we used the 8 nearest neighbors.

We applied the method with the smoothness term to the four image sequences; *hands*, *paper*, *back* and *heart*. The results are shown in Figure 1. The reconstruction error, both with and without the pairwise term, are reported in Table 2 for the cases when the ground truth is known. Note that while the labels are more visually pleasing the reconstruction quality is very similar. The problem sizes and parameters used for the experiments are shown in Table 1.

	r_0	r_k	λ	μ	Size of M
hands	5	3	5000	100	882×7899
paper	7	3	500	100	140×340
back	8	3	500	100	300×20561
heart	8	$\{3, 4\}$	500	100	160×68295

Table 1: Parameters used for the real image sequences. The column r_k contains the dimensions used for the local subspaces.

	Observed data			Missing data		
	Paper	Back	Heart	Paper	Back	Heart
$\mu = 0$	**1.07e2**	**6.86e2**	**1.50e3**	**2.86e2**	**1.44e3**	**3.67e3**
$\mu \neq 0$	1.50e2	7.61e2	1.79e3	3.63e2	**1.44e3**	3.70e3

Table 2: Comparison of reconstruction error for the three datasets where ground truth is available.

References

[1] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *Trans. Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001. 1
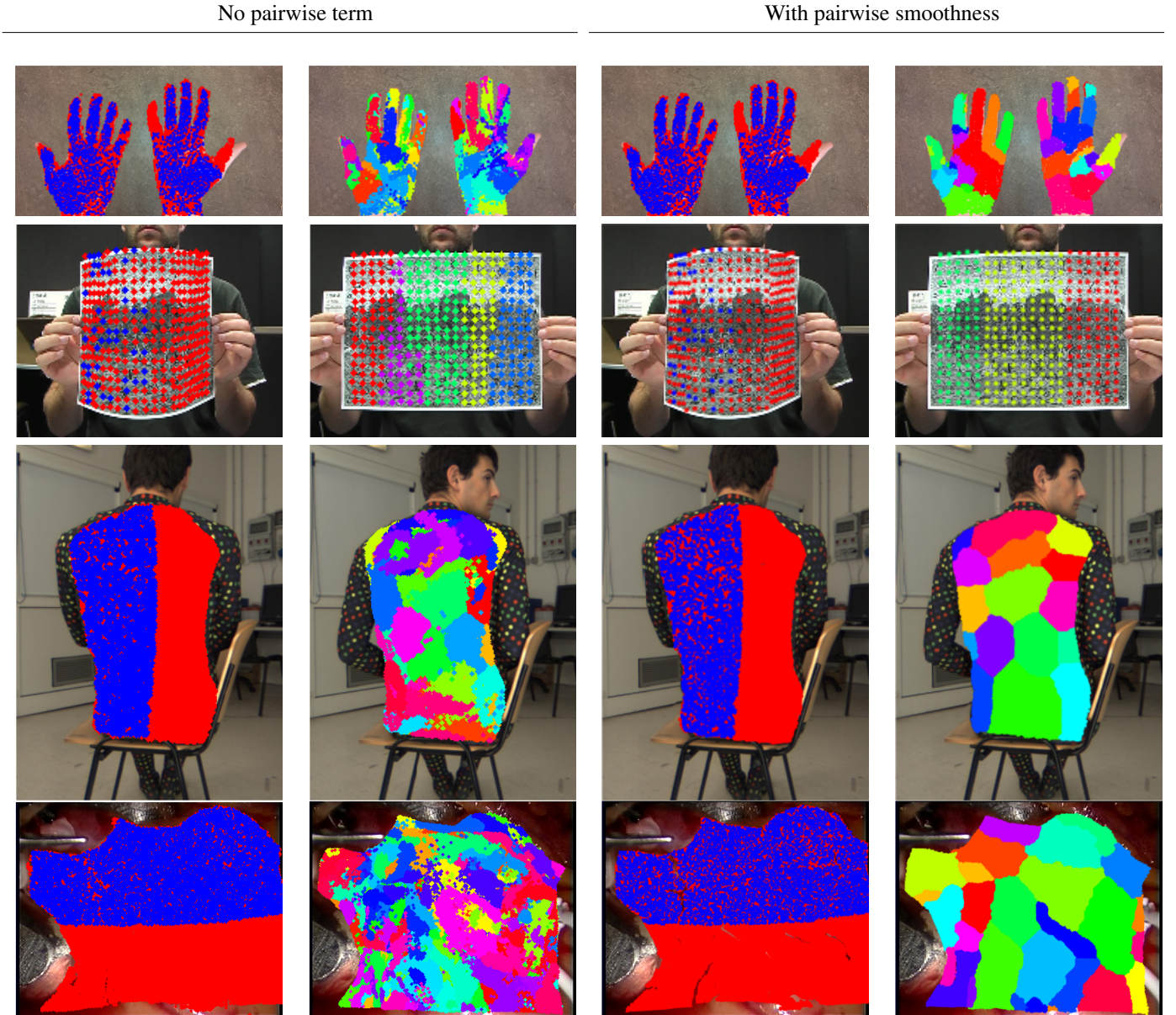
No pairwise term

With pairwise smoothness

Figure 1: Estimated point positions and obtained clusters using the energy (1) with and without the smoothness term S_{pq}.