DTF-Chinese

April 12, 2013

Abbreviation:	dtf-chinesechar
Number of instances:	100
Number of variables:	4992–17856 (median 8920)
Number of labels:	2
Number of factors:	16–64 times the number of variables
Order:	2
Function type:	explicit

Description The problem addressed is that of inpainting occluded regions of an image of a handwritten Chinese character [1]. The problem is an artificial task and was constructed so that higher-order structure between binary variables needs to be modelled in order to successfully solve the task.

The character originate from the KAIST Hanja2 database and were originally extracted from student admission forms. We selected random disjoint subsets of 300 training and 100 test characters; almost all characters are disjoint between the sets in that they appear in only one of the sets.

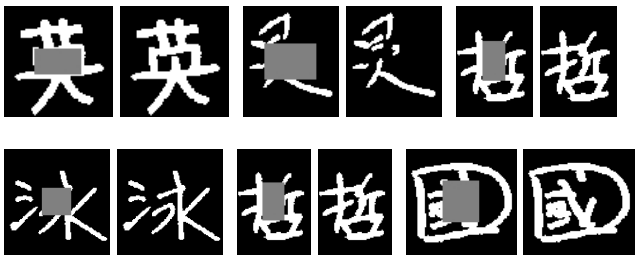

Figure 1: Input and ground truth output for a number of test instances. The input image consists of only three possible input intensities: $\{0, 127, 255\}$.

Objective / Learning We use a set of *factor types* $T = T_1 \cup T_2$ that instantiate a replicated set of interactions φ_t around each variable. Each interaction is conditioned on the image content I and the image position i. This yields the following energy.

$$J(x) = \sum_{t \in T_1} \sum_{i \in V} \varphi_t(x_i; I, i) + \sum_{t \in T_2} \sum_{i \in V} \varphi_t(x_i, x_{j(t)}; I, i), \tag{1}$$

As set T we use an 8-neighborhood at one and two pixels distance away, as well as a set of 27 neighbors at relative offsets $(-9, 0)$, $(-9, 3)$, $(-9, 6)$, $(-9, 9)$, $(-6, 0)$, $\ldots$, $(9, 9)$.

For each factor type the energies $\varphi_t(\cdot; I, i)$ are represented explicitly using learned decision trees. All parameters in the model are learned from a fully-observed training set using the maximum pseudolikelihood estimator from 300 training images. Because the energies depend on the local image content, the resulting field is heterogeneous.

The learned energies are interesting in that short range interactions tend to be submodular, whereas longer range interactions tend to be supermodular. The mix of these two types of interactions at different ranges yields challenging energy minimization problems.

References

[1] Sebastian Nowozin, Carsten Rother, Shai Bagon, Toby Sharp, Bangpeng Yao, and Pushmeet Kohli. Decision tree fields. In *ICCV*, pages 1668–1675. IEEE, 2011.