## Supplementary Material: Efficient Label Tree Learning For Object Recognition

Jia Deng<sup>1,2</sup>, Sanjeev Satheesh<sup>1</sup>, Alexander C. Berg<sup>3</sup>, Li Fei-Fei<sup>1</sup> Computer Science Department, Stanford University<sup>1</sup> Computer Science Department, Princeton University<sup>2</sup> Computer Science Department, Stony Brook University<sup>3</sup>

## **1 Proofs**

Lemma 1.1. For LP problem

 $\begin{array}{ll} \underset{x}{\mininize} & c^T x\\ subject \ to & Ax \leq b\\ & 0 \leq x \leq \end{array}$ 

where  $A \in \mathbb{R}^{m \times n}$ , m < n, if it is feasible, then there exists an optimal solution with at most m non-integer entries and such a solution can be found in polynomial time.

*Proof.* Let  $x^*$  be an optimal solution that an LP solver returns. Let B be the set of indices of entries in  $x^*$  that are non-integers,  $B = \{i : x_i^* \in (0, 1)\}$  and let E be the rest of the indices.

If  $|B| \le m$ , then we are done. We now consider the case when |B| > m.

Let *H* be the polyhedron  $H = \{x_B : c_B^T x_B = c^T x_B^*, A_B x_B = A_B x_B^*, 0 \le x_B \le 1\}$ , where  $A_B$  is the columns indexed by *B*. Observe that any *x* such that  $x_B \in H$  and  $x_E = x_E^*$ , is an also an optimal solution of the LP. That is, replacing the non-integer entries of  $x^*$  with those in *H* still gives an optimal solution.

Since  $x_B^* \in H$ , therefore H is non-empty. Also H is bounded. Hence there exists at least one *basic* feasible solution  $x_B'$  of H(Bertsimas & Tsitsiklis [1]), for which there are |B| linearly independent constraints that are active. Such a basic feasible solution can be found in polynomial time by solving an auxiliary LP by introducing additional artificial variables, the same as the Phase 1 of the simplex method. Details can be found in [1].

We now show that  $x'_B$  has at most m non-integer entries.

We first show that  $\forall x_B \in \text{null}(A_B)$ ,  $c_B^T x_B = 0$ . Assume to the contrary that there exists  $\hat{x}_B \in \text{null}(A_B)$  such that  $c_B^T \hat{x}_B < 0$ . Let  $y^* \in \mathbb{R}^n$  be such that  $y_B^* = x_B^* + \theta \hat{x}_B$  and  $y_E^* = x_E^*$ , where  $\theta > 0$ . It follows that for sufficiently small  $\theta$ ,  $y^*$  satisfies all contraints of the LP, since  $Ay^* = Ax^* + \theta A_B \hat{x}_B = Ax^* \leq b$  and  $0 \leq y_B^* = x_B^* + \theta \hat{x}_B \leq 1, 0 \leq y_E^* = x_E^* \leq 1$ . Also the LP has a smaller value, since  $c^T y^* = c^T x^* + \theta c_B^T \hat{x}_B < c^T x^*$ , which is contradition.

If follows that  $c_B \in \operatorname{null}(A_B)^{\perp} = \operatorname{row}(A_B)$ . Therefore the number of linearly independent vectors among  $c_B$  and rows of  $A_B$  is at most m. Since  $x'_B$  has |B| > m linearly independent constraints that are active, at least |B| - m constraints from  $0 \le x'_B \le 1$  must be active and therefore at least |B| - m entries of  $x'_B$  are integers. Hence  $x'_B$  has at most m non-integer entries.

We then replace the entries  $x_B^*$  in  $x^*$  with  $x_B'$  and obtain an optimal solution with at most m non-integer entries.

## References

[1] D. Bertsimas and J.N. Tsitsiklis. Introduction to linear optimization. 1997. 1