Implementation of BP for Early Vision

Pedro F. Felzenszwalb and Daniel P. Huttenlocher

February 13, 2005

This document briefly describes an implementation of the algorithms from our paper **Efficient Belief Propagation for Early Vision** which appeared in CVPR 2004.

Please email comments and bug reports to: *pff@cs.uchicago.edu*.

1 Usage

- Compile the programs by running "make".
- Example usage of the restoration program:
  - run "./noise penguin.pgm noisy.pgm 20" to create a noisy image.
  - the noisy image is saved in "noisy.pgm".
  - run "./restore noisy.pgm result.pgm" to restore the noisy image.
  - the restored image is saved in "result.pgm".
- Example usage of the stereo program:
  - run "./stereo tsukuba1.pgm tsukuba2.pgm result.pgm".
  - the disparity map is saved in "result.pgm".

2 Parameters

The general framework for the problems we consider is as follows (see the paper for details). Let $\mathcal{P}$ be the set of pixels in an image and $\mathcal{L}$ be a set of labels. A labeling f assigns a label $f_p \in \mathcal{L}$ to each pixel $p \in \mathcal{P}$. The quality of a labeling is given by an energy function:

$$E(f) = \sum_{(p,q)\in\mathcal{N}} V(f_p, f_q) + \sum_{p\in\mathcal{P}} D_p(f_p),$$

where $\mathcal{N}$ are the edges in the four-connected image grid graph, $V(f_p, f_q)$ is the cost of assigning labels f_p and f_q to two neighboring pixels, and $D_p(f_p)$ is the cost of assigning label f_p to pixel p.

2.1 Image Restoration

For image restoration the labels are gray levels. The file "restore.cpp" implements the image restoration algorithm using the following parameters for the energy function:

$$\begin{aligned} V(f_p, f_q) &= \min\left((f_p - f_q)^2, \text{DISC\_K}\right), \\ D_p(f_p) &= \lambda \min\left((I(p) - f_p)^2, \text{DATA\_K}\right). \end{aligned}$$

Where DISC_K, DATA_K and λ are compile time defines.

2.2 Stereo

For stereo the labels are disparities. The file "stereo.cpp" implements the stereo algorithm using the following parameters for the energy function:

$$\begin{aligned} V(f_p, f_q) &= \min\left(|f_p - f_q|, \text{DISC\_K}\right), \\ D_p(f_p) &= \lambda \min\left(|I_l(x, y) - I_r(x - f_p, y)|, \text{DATA\_K}\right). \end{aligned}$$

Where DISC_K, DATA_K and λ are compile time defines. The stereo program also has a few other important compile time parameters: VALUES is the maximum disparity plus one, SCALE controls how disparities are scaled to graylevels in the output image, and SIGMA is the amount to smooth the input images before processing.