Homework #2: Similarity Measure

Due February 26th, 2014
Professor Davi Geiger.

The goal of this assignment is to develop a measure of similarity for stereo matching. The data set is from http://vision.middlebury.edu/stereo/data/2014/

**Morlet Wavelets:** You have done all of this already, except, let us use fewer angles, only 6, not 8, to reduce computations. See below

The Morlet wavelet used is

\[ \psi_{\sigma, \theta}(u) = \frac{C_1}{\sigma} \left( e^{i \frac{\pi}{2\sigma}(u \cdot e_\theta)} - C_2 \right) e^{-\frac{u^2}{2\sigma^2}} \]

where \( e_\theta = (\cos \theta, \sin \theta) \), with parameters \( \sigma, \theta \):

(a) \( \sigma = 1 \), (b) \( \sigma = 3 \), (c) \( \sigma = 6 \),

(i) \( \theta = 0 \), (ii) \( \theta = \frac{\pi}{6} \), (iii) \( \theta = \frac{\pi}{3} \), (iv) \( \theta = \frac{\pi}{2} \), (v) \( \theta = \frac{2\pi}{3} \), (vi) \( \theta = \frac{5\pi}{6} \),

Let us refer to the parameters as \( \lambda = (\sigma, \theta) \), i.e., there are \( \lambda_1, \ldots, \lambda_{18} \) different values.

A number and two vector of 18 entries (numbers) each, per pixel \( u \).

Normalize the input images (make sure they are double) by dividing its values by 255. Convolve the image with a Gaussian Blur

\[ S^0 I(u) = G_{\sigma=6} * I(u) = C e^{-\frac{u^2}{2\sigma^2}} * I(u) = \sum_{x'} \sum_{y'} I(x-x', y-y') C e^{-\frac{(x')^2 + (y')^2}{2\sigma^2}} \]

This is one number per pixel. Let us save it.

For each image and each \( \lambda_i = (\sigma, \theta) \), perform a convolution with the real part of the wavelet and separately with the complex part and produce the results,

\[ W_{\lambda_i} I(u) = \psi_{\lambda_i} * I(u) = \sum_{x'} \sum_{y'} I(x-x', y-y') \psi_{\lambda_i}(x', y') \]

or

\[ \psi_{\lambda_i}(x', y') \]
\[ W_{\lambda_i}^{\text{Real}} I(u) = \psi_{\lambda_i}^{\text{Real}} * I(u) = \sum_{x'} \sum_{y'} I(x - x', y - y') \psi_{\lambda_i}^{\text{Real}}(x', y') \quad i = 1, \ldots, 18 \]

\[ W_{\lambda_i}^{\text{Im}} I(u) = \psi_{\lambda_i}^{\text{Im}} * I(u) = \sum_{x'} \sum_{y'} I(x - x', y - y') \psi_{\lambda_i}^{\text{Im}}(x', y') \quad i = 1, \ldots, 18 \]

Store \( v^{\text{Real}} I(u) = \{ W_{\lambda_i}^{\text{Real}} I(u); i = 1, \ldots, 18 \} \) as a vector at each pixel \( u \) and \( v^{\text{Im}} I(u) = \{ W_{\lambda_i}^{\text{Im}} I(u); i = 1, \ldots, 18 \} \) as another vector at each pixel \( u \).

So all together, we have at each pixel \( u \), a number \( S^0 I(u) \) and two vectors \( v^{\text{Real}} I(u) \), \( v^{\text{Im}} I(u) \) of length 18.

Let us represent all together as \( f(u) = \{ S^0 I(u), v^{\text{Real}} I(u), v^{\text{Im}} I(u) \} \).

**Problem 1: Detect Sparse Features**

Given threshold values \( (T_h^{\text{Real}}, T_h^{\text{Im}}) \), let us only consider locations \( u \) where they satisfy

\[ |v^{\text{Real}} I(u)|^2 = \sum_{i=1}^{18} (W_{\lambda_i}^{\text{Real}} I(u))^2 < T_h^{\text{Real}} \]

\[ |v^{\text{Im}} I(u)|^2 = \sum_{i=1}^{18} (W_{\lambda_i}^{\text{Im}} I(u))^2 > T_h^{\text{Im}} \]

If there are nearby pixels that pass the detection, select a location \( u \) such that \( |v^{\text{Im}} I(u)|^2 \) is maximum. How to choose \( (T_h^{\text{Real}}, T_h^{\text{Im}}) \)? Make sure that no more than 1% of the image pixels pass the test. So lower \( T_h^{\text{Real}} \) and raise \( T_h^{\text{Im}} \) so that you end up with less than 1% of the number of pixels. Apply this method to the Left image, choosing the threshold values. Then, using the same threshold values, apply the method to the Right image.

**Problem 2: A Similarity Measure**

Given a stereo pair of images and a pixel location \( I^L(u) \) and \( I^R(u) \), we define a similarity measure as a number that matches pixels \( u \) and \( v \) on the left and right images. We consider for each left image, a template around pixel \( u = (x, y) \). Let us consider
templates $T\{(x - N_x; x + N_x, y - N_y; y + N_y)\}$, i.e., $T$ has size $(2 * N_x + 1, 2 * N_y + 1)$. Let us choose $N_x = N_y = 16$.

Consider a disparity $(d_u, 0)$ at pixel $u = (x, y)$. The similarity measure is given by

$$
\text{Similarity} \left( I^L(u), I^R(u - (d_u, 0)) \right) = \frac{1}{(2 * N_x + 1) * (2 * N_y + 1) *}
$$

$$
\sum_{y' = y - N_y}^{y + N_y} \sum_{x' = x - N_x}^{x + N_x} \left[ (S^0 I^L(x', y') - S^0 I^R(x' - d_u, y'))^2 
+ \frac{1}{18} \left( v^{\text{Real}} I^L(x', y') - v^{\text{Real}} I^R(x' - d_u, y') \right)^2 
+ \frac{1}{18} \left( v^{\text{Im}} I^L(x', y') - v^{\text{Im}} I^R(x' - d_u, y') \right)^2 \right]
$$

For all pixels $u$ in the Left Image that passed the threshold values test, question 1, consult the true disparity provided in the data set website, http://vision.middlebury.edu/stereo/data/2014/ call it $d_u^{\text{true}}$. Compute $\text{Similarity} \left( I^L(u), I^R(u - (d_u^{\text{true}}, 0)) \right)$.

2a. Create a histogram: The x-axis is the Similarity values. The y-axis represents the counting. For every pixel $u$ place one vote at x-axis value $\text{Similarity} \left( I^L(u), I^R(u - (d_u^{\text{true}}, 0)) \right)$. Note that “bins” must be created on the x-axis so that similar values of $\text{Similarity} \left( I^L(u), I^R(u - (d_u^{\text{true}}, 0)) \right)$ are voted to the same bin. We must find the bin size to be large enough so that many votes hit the same bin, and not too large to avoid that all votes hit the same bin. Producing a distribution over the bins is the desired result.

2b. Create a second histogram. This time, for every pixel $u$ in the Left Image that that passed the threshold values test, question 1, and the correspondent $d_u^{\text{true}}$, find two pixels $v$ in the Right Image of the form $v = u - (d_u^{\text{true}} + \Delta d, 0)$, with $\Delta d \neq 0$, that passed the threshold values test of question 1. Try both neighbors, for $\Delta d < 0$, and $\Delta d > 0$, and the neighbors are characterized by $|\Delta d|$ as small as possible (not zero). Create a similar histogram as 2a, with the same bin size, but with these similarity values.

2c. By comparing the histograms in 2a and 2b we can create a threshold such that it is high enough so that most matches in histogram 2a would pass and low enough so that most matches in 2b would not pass.