We will first study the array

\[ C_{e}(x+w) = \phi_{\text{intensity}}(e,x,w,\theta = 0) + \phi_{\text{intensity}}(e,x,w,\theta = \pi/2) + \phi_{\text{contrast}}(e,x,w,\theta = 0) + \phi_{\text{contrast}}(e,x,w,\theta = \pi/2) \]

\[ \phi_{\text{intensity}}(e,x,w,\theta) = 1 + 0.05 \frac{W_{e}(x)}{T} \]

\[ \phi_{\text{contrast}}(e,x,w,\theta) = \frac{\left| D_{e}^{-}(e,x,w,s) = 3 \right| + 1}{\left| D_{e}^{+}(e,x,w,s) = 3 \right|}^{p} + 1 \]

where \( p = 0.7 \)

defined for each epipolar line and limited to a band \(|w| \leq D\). We will study the data at scale \( s = 3 \).

Note that \( l = x - w/2 \) and \( r = x + w/2 \). For \( x + w \) odd, \( l \) and \( r \) are at subpixel locations. In these cases, both the intensity and the derivatives are then defined as the average values for the neighbor integer values of \( l \) and \( r \) respectively. More precisely,

\[ \hat{I}_{e}(l,\theta,s) = \frac{1}{2} \left( \hat{I}_{e}(l,\theta,s) + \hat{I}_{e}(l+1,\theta,s) \right) \quad \text{and} \quad \hat{I}_{r}(e,r,\theta,s) = \frac{1}{2} \left( \hat{I}_{r}(e,r,\theta,s) + \hat{I}_{r}(e,r+1,\theta,s) \right) \]

\[ D\hat{I}_{e}(l,\theta,s) = \frac{1}{2} \left( D\hat{I}_{e}(l,\theta,s) + D\hat{I}_{e}(l+1,\theta,s) \right) \quad \text{and} \quad D\hat{I}_{r}(e,r,\theta,s) = \frac{1}{2} \left( D\hat{I}_{r}(e,r,\theta,s) + D\hat{I}_{r}(e,r+1,\theta,s) \right) \]

\[ D_{e}^{-}(e,x,w) = \left| D\hat{I}_{r}(e,r,\theta,3) + D\hat{I}_{e}(l,\theta,3) \right|, \quad D_{e}^{+}(e,x,w) = \left| D\hat{I}_{r}(e,r,\theta,3) - D\hat{I}_{e}(l,\theta,3) \right| \]

1. Synthetic two “color” (B & W) images (for debugging purposes)

For epipolar line \( e = 32 \) and \( D = 6 \) and \( s = 3 \), display \( C[x,w+D] \) as an image \( I[x,w+D] \) of length \( 2N \) and height \( 2D+1 \), where the grey values represent the values of \( C \) along the selected epipolar line. To avoid boundary problems, examine from \( x = D, \ldots, 2N-1-D \). In order to enhance, discriminate, and examine the regions where the matches are good, i.e., \( C[x,w+D] \) is small, we apply a log function and normalize the results to vary from 0 to 255. We use the formula

\[ \bar{C}[x,w+D] = \log \left( 1 + C[x,w+D] \right) \]

\[ I[x,w+D] = \text{floor} \left[ 255 \frac{(\bar{C}[x,w+D] - \bar{C}_{\min})}{\bar{C}_{\max} - \bar{C}_{\min}} \right] \quad \text{with} \quad x = 6, \ldots, 2N-7 \quad \text{and} \quad w = -6, \ldots, 6 \]

where \( \bar{C}_{\max} = \max_{x,w} \bar{C}[x,w+D] \), \( \bar{C}_{\min} = \min_{x,w} \bar{C}[x,w+D] \)
Note that the image $I[x,w+D]$ is twice as wide as the original image, since $x=1,...,2N$, and much smaller in height, since $w+D = 0,...,2D$.

Example of values of $I[x,w+D]$ for the random dot stereogram.

Example of values of $I[x,w+D]$ for the square stereo.

Display 2 different images for each stereo pair, with two different values of $T=20, 120$

2. **Full synthetic images.** Now we consider the entire image, i.e., we examine every epipolar lines. Display an array for the non-negative value function $\tilde{d}(x,e) = w^*(x,e) + D$, of length $2N$ and width $N$, where

$$w^*(x,e) = \arg\min_{w \in (-D,..., +D)} C[x,w+D]$$

for $x=6,...,2N-6$ and $e=s,...,N-s$

is the “best” disparity value, the one that minimizes $C[x,w+D]$, and

$$\tilde{d}(x,e) = w^*(x,e) + D$$

is a non-negative value function representing $w^*(x,e)$

Note: at many locations, $(x,e)$, there will be ambiguities, i.e., many different disparities $w$ will give the same (maximum) value $C[x,w+D]$. In this case, either take the last or first disparity value (so no need to modify the code) or use the smallest disparity value, but be consistent.

In order to display $\tilde{d}(x,e)$ we scale it up to 256 values through the linear formula

$$I[x,e] = \text{floor}\left[ 255 \frac{\tilde{d}(x,e) - \tilde{d}_{\min}}{\tilde{d}_{\max} - \tilde{d}_{\min}} \right]$$

where $\tilde{d}_{\max} = \max_{x,e} \tilde{d}(x,e)$, $\tilde{d}_{\min} = \min_{x,e} \tilde{d}(x,e)$

Note that the resulting array is twice as wide as the original left and right image, since $x=1,...,2N$, and has the same height $e=1,...,N$. The explanation is that we are showing subpixel disparity results as if it was a pixel. For the random dot stereogram, the display of the array $I[x,e]$ should look like this (grey values are a multiplication factor of the disparity values). The results at this stage are not expected to be as clean/good as a full stereo algorithm (next homework).
Example of $I[x,e]$ for (a) the random dot stereogram and (b) the square. Note the horizontal stretch since there are twice as many pixels (sub pixels) along $x$.

3. **Two Grey scale stereo pairs.** Now the disparity may vary more, for the “pentagon” image pair use $D_{max}=D=15$, and for the “tsukuba” image pair use $D_{max}=D=65$. Also, in real images, one may be darker than the other (the aperture of the camera may not be the same). So the first step, before applying any filter, is to obtain the mean value of all pixels in the left and right images, $\mu_L$ and $\mu_R$, and add to every pixel in the darker image the difference $|\mu_L-\mu_R|$, so that the mean value of the final two images is the same. This is particularly helpful in the “pentagon” image.

Display an array of length $2N$ and width $N$ with non-negative values $ar{d}(x,e) = d^*(x,e) + D$, where

$$d^*(x,e) = \arg\min_{w \in (D, \ldots, +D)} C[x,w+D], \text{ for } D, \ldots, 2N-D \text{ and } e=s, \ldots, N-s$$

So $d^*(x,e)$ is the value of the “best” disparity value, the one that maximizes $C[x,w+D]$

$$I[x,e] = \text{floor} \left[ 255 \left( \frac{\bar{d}(x,e) - \bar{d}_{\text{min}}}{\bar{d}_{\text{max}} - \bar{d}_{\text{min}}} \right) \right]$$

where $\bar{d}_{\text{max}} = \max_{x,e} \bar{d}(x,e)$, $\bar{d}_{\text{min}} = \min_{x,e} \bar{d}(x,e)$

Note that the resulting array is twice as wide as the original left and right image, since $x=1, \ldots, 2N$, and has the same height $e=1, \ldots, N$. 
The values of $I[x,e]$ for the pentagon image

4. Stereo Algorithm

The algorithm runs for each epipolar line separately. We create two arrays $F_1$ and $F_2$ that will hold all the transition costs

$$F_1[x, w + D, w' + D] = TiltCost[w - w'] \quad |w' - w| \leq 1$$

$$F_2[x, w + D, w' + D] = \begin{cases} \text{Occlusion} & \phi_{\text{contrast}}(e, x, w, \theta = 0) \\ \text{Occlusion} & \phi_{\text{contrast}}(e, x - 1 - |w - w'|, w', \theta = 0) \end{cases} \quad w' \geq w + 1$$

$$C_1[x, w + D] = \phi_{\text{intensity}}(e, x, w, \theta = 0) + \phi_{\text{intensity}}(e, x, w, \theta = \pi/2) + \phi_{\text{contrast}}(e, x, w, \theta = 0) + \phi_{\text{contrast}}(e, x, w, \theta = \pi/2)$$

We then use the arrays $F_1[x, w+D, w'+D]$ , $F_2[x, w+D, w'+D]$ and $C[x, w+D]$ to compute the best disparity map. Note that when programming $F_1[x, w+D, w'+D]$ may be more compactly written if $w'+D \rightarrow w+D+i$ where $i=-1,0,1$. Then $F_1[x, w+D, w'+D] \rightarrow F_1[x, w+D, w'+1], i=0,1$. The parameters $TiltCost$ and $Occlusion$ will be estimated through testing. Let us apply dynamic programming line by line, i.e., epipolar line by epipolar line. Define disparity as $d^*(x,e)=w^*(x,e)+D$.

For each stereo pair display the “Disparity” array

$$\text{Disparity}[x,e] = \text{floor} \left[ 255 \frac{d^*(x,e) - \overline{d}_{\text{min}}}{\overline{d}_{\text{max}} - \overline{d}_{\text{min}}} \right]$$

where $\overline{d}_{\text{max}} = \max_{x,e} d^*(x,e)$ , $\overline{d}_{\text{min}} = \min_{x,e} d^*(x,e)$

Below is a dynamic programming algorithm to obtain $d^*(x,w)$, i.e., to obtain $w^*(x,e)$. 
**Stereo-Matching DP** (ImageLeft, ImageRight, D, e) (e-is epipolar line)

**Initialize**
/* precomputation of the match cost and transition cost, stored in array F */

**loop for** v=(x,w) in V (i.e., loop for x and loop for w)

Set-Array C[x,w+D]

**loop for** w'=-D,…,0,…D

Set-Array F1[x, w+D, w'+D] and F2[x, w+D, w'+D]

end loop

end loop

**Main loop**

**loop for** x=D, …, 2N-1-D

**loop for** w=-D,….0,…,D

Cost=∞;

? **loop for** w’=-D,…,0,…,D

Dw’=|w-w’|

if (Dw’ ≤ 1) {

Temp= Φ x-1*[w’+D]+ F1[x, w+D, w’+D];

if (Temp < Cost) {

Cost= Temp;

back[x,w+D] =w’+D;

bestw=2D+10;

}

if (Dw’ ≠ 0) { /* w=w’+1 and w=w’-1 are being considered again */

x’=x-1; Dw’;

Temp= Φ x’*[w’+D]+ Dw’*F2[x, w+D, w’+D];

if (Temp < Cost) {

Cost= Temp;

bestw=w’;

}

end loop

if (bestw < D+1) { w’=bestw;

if (w > w’)

**loop for** i=0… Dw’ back[x-i,w+D] =w+D;

if (w < w’) { back[x,w+D]=w’+D

**loop for** i=1 … Dw’ back[x-i,w’+D]=w’+D

Φ,x*[w+D]= Cost+ C[x,w+D];

end loop

end loop

**Backtrack**

x=2N-1-D;

min=∞;?

**loop for** w=-D,….0,…,D

if (Φ,x*[w+D] < min) {

min= Φ,x*[w+D];

d*[x,e]=w+D;

}

**loop for** x=2N-1-D, ….D+1

d*[x-1,e]= back[x, d*[x,e]] ;
Expected disparity results for the random dot stereogram and for the pentagon stereo pair.