Affine Transform Based Image Rectification For Better Disparity From Stereo Matching

Sunil Kumar Kopparapu
SunilKumar.Kopparapu@TCS.Com

Cognitive Systems Research Laboratory
Tata Consultancy Services Limited, Navi Mumbai

http://www.tcs.com

November 2006

Disparity from Stereo Images

Let $L(x,y)$ and $R(x,y)$ be the image captured by the left camera and the right camera respectively.

Say, $L(x_l,y_l)$ in the left image and $R(x_r,y_r)$ in the right image correspond to the same object in the 3D scene. Then,

Disparity $\mathcal{D}(x,y) \overset{\text{def}}{=} (\mathcal{D}(x),\mathcal{D}(y))$,
where $\mathcal{D}(x) = x_l - x_r$ and $\mathcal{D}(y) = y_l - y_r$

Disparity is the *spatial differences* between the corresponding points in the 3D scene, captured by the left and the right camera.

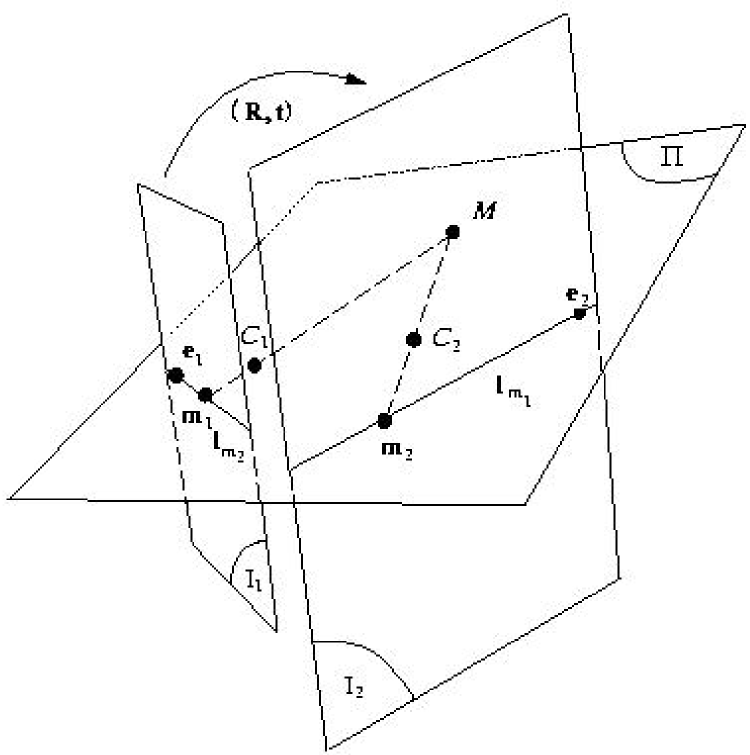
An Example - Stereo Image; Disparity

Left Image (L) Right Image (R) Disparity ($\mathcal{D}$)

Note: Darker region in $\mathcal{D}$ means region in 3D scene further from camera*.

*http://www-dbv.cs.uni-bonn.de/stereo_data/

Epipolar Aligned Images



C_1 and C_2 are the two cameras;
I_1 and I_2 are the image planes;
M is the object in 3D scene;
m_1 and m_2 are the projections of M;
II is the epipolar plane;
l_{m_1} and l_{m_2} are the epipolar lines

Note: Most stereo algorithms assume images to be epipolar aligned, namely $\mathcal{D}(y) = 0$. **Why?**

Advantage: Epipolar Aligned Images

- Disparity estimation is a 2D search problem.

- Epipolar aligned

  - $\mathcal{D}(y) = 0 \Rightarrow y_l = y_r$

  - *spatial differences* between the images in only in one direction (the x-direction)

- the 2D search disparity estimation problem reduces to a 1-D search problem.

Why Stereo Image Rectification

Image rectification loosely defined as a process of rectifying the image to take care of geometrical distortion creeping into the imaging system.

Image rectification in stereo images would essentially mean epipolar alignment of the images

Note: For an off the shelf CCD camera with a lens of 8 mm focal length and pixel size 11 μm, an error of 1^o about the optical axis results in a shift of 0.16 mm ($\tan(1^o) \times 8$).

$\Rightarrow$ 14 ($0.16 \times 10^{-3}/11 \times 10^{-6}$) pixels shift across the line (y - axis). Violates epipolar line constraint and is for most stereo vision algorithms unacceptable.

Approach: Stereo Image Rectification - 1

Assume that we are able to identify distinct points corresponding to the same scene point in the two stereo images *fairly* accurately.

Let $\{(x_l^1, y_l^1), (x_l^2, y_l^2), \cdots (x_l^N, y_l^N)\}$; $\{(x_r^1, y_r^1), (x_r^2, y_r^2), \cdots (x_r^N, y_r^N)\}$ be the corresponding points in the left and the right image

Geometrically rectify the images so that the same point in the scene have the same y-ordinate in the left and the right images.

How?

Approach: Stereo Image Rectification - 2

Consider the following relation between left and the right image

$$y_l^i = a_0 + a_1 x_r^i + a_2 y_r^i \quad \text{for} \quad i = 1, 2, \cdots N$$

This can be written in the matrix form as $L = RA$ where,

$$L = \left[y_l^1 \ y_l^2 \ \cdots \ y_l^N \right]^T; \quad A = [a_0 \ a_1 \ a_2]^T; R = \begin{bmatrix} 1 & y_r^1 & x_r^1 \\ 1 & y_r^2 & x_r^2 \\ \vdots & \vdots & \vdots \\ 1 & y_r^N & x_r^N \end{bmatrix}$$

Need to solve $\min_A \|L - RA\|_2^2$ for A. The minimum least squares norm solution

The least square norm solution is given by

$$A = (R^T R)^{-1} R^T L$$

Stereo Image Rectification - Steps

```
for i=1 to rows do
    for j=1 to cols do
        x̂_r(i,j) = x_r(i,j)
        ŷ_r(i,j) = a_0 + a_1 y_r(i,j) + a_2 x_r(i,j)
        {a_0, a_1, a_2 are obtained solving matrix equation.}
        R̂(i,j) = R(x̂_r(i,j), ŷ_r(i,j))
        {R is the right image and the R̂ is the rectified image }
    end for
end for
```

for i=1 to rows **do**

  for j=1 to cols **do**

    $\hat{x}_r(i,j) = x_r(i,j)$

    $\hat{y}_r(i,j) = a_0 + a_1 y_r(i,j) + a_2 x_r(i,j)$

    $\{a_0, a_1, a_2$ are obtained solving matrix equation.$\}$

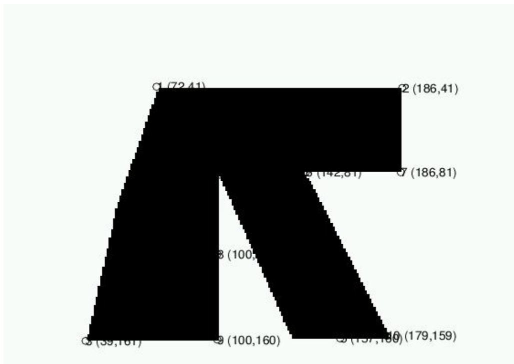
    $\hat{R}(i,j) = R(\hat{x}_r(i,j), \hat{y}_r(i,j))$

    $\{R$ is the right image and the $\hat{R}$ is the rectified image $\}$

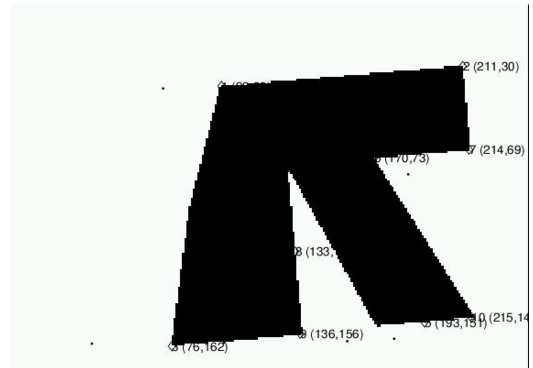
  end for

end for

Stereo Image Rectification - Process

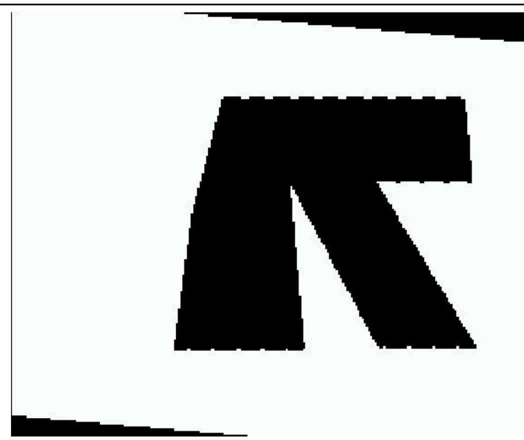


Left Image



Right Image

Rectification Process



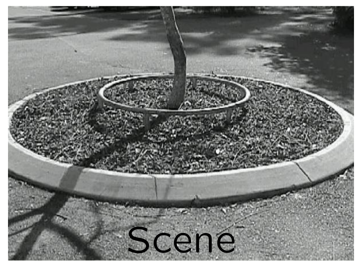
Rectified Image

Affordable Stereo Camera - 1


*

Left Camera

Right Camera



Scene

Left Image

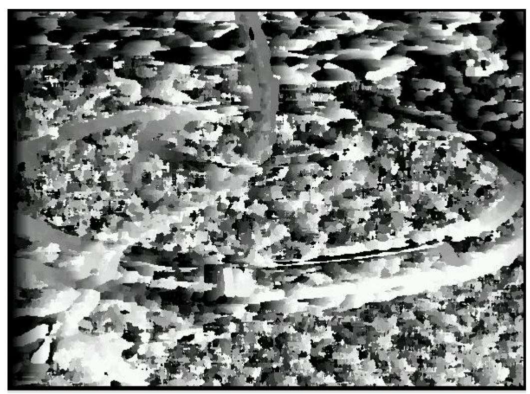
Right Image

*http://en.wikipedia.org/wiki/Digital_camera

Affordable Stereo Camera - 2

Left Image

Right Image

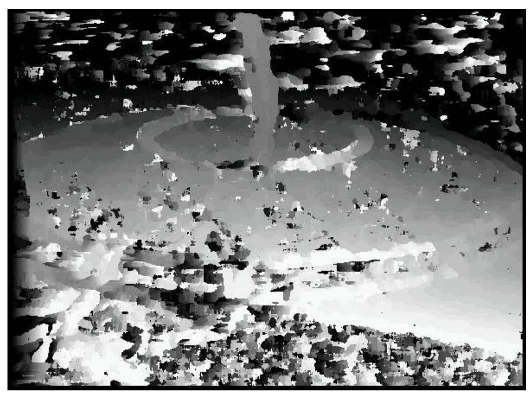


Disparity Map without Rectification

Affordable Stereo Camera - 3

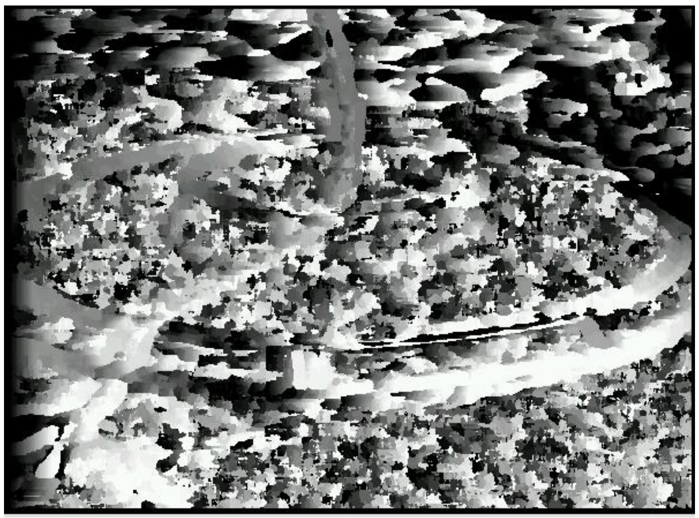
Left Image

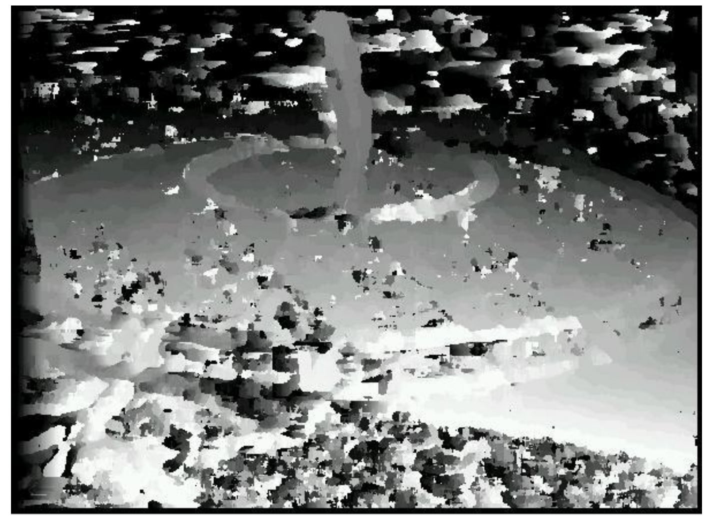
Rectified Right Image



Disparity Map after rectification

Affordable Stereo Camera - 4





Disparity map before and after epipolar alignment

Conclusions

- a simple scheme to rectify stereo image pair to produce epipolar aligned images
- can be used as a preprocessor to stereo algorithms
- the scheme can help in getting a significantly improved disparity map
- use a hand held camera to generate stereo images (affordable stereo camera).