We present a method for calculation of disparity maps from stereo sequences. Disparity map from previous frame is first transferred to the new frame using estimated motion of the calibrated stereo rig. The predicted disparities are validated for the new frame and areas where prediction failed are matched with a traditional stereo matching algorithm. This method produces very fast and temporally stable stereo matching suitable for real-time applications even on non-parallel hardware.

**Main Processing Cycle**

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{L,k}$</td>
<td>$k$-th frame left image</td>
</tr>
<tr>
<td>$I_{R,k}$</td>
<td>$k$-th frame right image</td>
</tr>
<tr>
<td>$I_{L,k-1}$</td>
<td>Previous left image</td>
</tr>
<tr>
<td>$D_k$</td>
<td>Disparity map for frame $k$</td>
</tr>
<tr>
<td>$D_{k-1}$</td>
<td>Previous disparity map</td>
</tr>
<tr>
<td>$\hat{D}_k$</td>
<td>Disparity map validation</td>
</tr>
<tr>
<td>$\hat{D}_k$</td>
<td>Disparity map prediction</td>
</tr>
<tr>
<td>$R_t$, $\hat{D}_k$</td>
<td>Ego-motion estimation</td>
</tr>
<tr>
<td>$\hat{D}<em>k$, $D</em>{k-1}$</td>
<td>Disparity map</td>
</tr>
<tr>
<td>$\hat{D}<em>k$, $D</em>{k-1}$</td>
<td>Disp. completion by seed growing</td>
</tr>
</tbody>
</table>

**Assumptions**

- Stereo rig calibration is known ($K, f, b$).
- Scene is static, camera may be moving.

**Algorithms**

- Disparity prediction (per pixel):
  \[
  x'_d = M \cdot x_d
  \]
  where $x_d = (x, y, w, d)$ consists of homogenous image coordinates $(x, y, w)$ and disparity $d$.

  \[
  M = \begin{bmatrix}
  KRK^{-1} \\
  0 \\
  T
  \end{bmatrix}
  \]

- Ego-motion estimation:
  Harris points, prediction by previous disparity map, Lucas-Kanade tracker and P3P in RANSAC with reprojection error.

- Disparity completion and bootstrapping:
  Seed growing stereo based on [1], seed queue re-insertion from predicted and validated disparity map.

**Typical Disparity Maps**

- Left
  - Per Frame Seed Growing (RTG)
  - Prediction + Growing (RTP)

- Right
  - Density of the disparity map
  - RMS error in traffic scene with ground truth
  - Error rate

**Comparison with Others**

- Compared with ELAS [2] and SGM [3]
- Evaluation on a synthesized sequence
- Error (per pixel) when $|d - d_{GT}| > 1$
- RMS error low for both RTP and RTG
- Challenging data: lower density, but stable error rate

**Real-time Implementation**

The computation is split into two threads working in a producer-consumer pattern:

- Initialization – precompute MNCC, Harris
- Stereo – ego-motion, prediction, growing

Running 20fps with resolution $640 \times 480$ on Dell Latitude E6420 Quad Core (using two cores).

Video capture with Bumblebee2 stereo camera (SDK used only for rectification).

**Conclusion**

- Resulting disparity maps contain less flicker
- Very low computational demands
- Not sensible to disparity search range
- Ego-motion estimate as a bonus
- Prediction failures are not fatal – handled by traditional stereo

**Acknowledgement**

This work was supported by the Czech Ministry of Education under Project MEB111006.

This work was presented at ICCV 2011 Workshop on Live Dense Reconstruction from Moving Cameras.

**Reference**