Integrating Stereo with Shape-from-Shading derived Orientation Information

Tom S. F. Haines and Richard C. Wilson

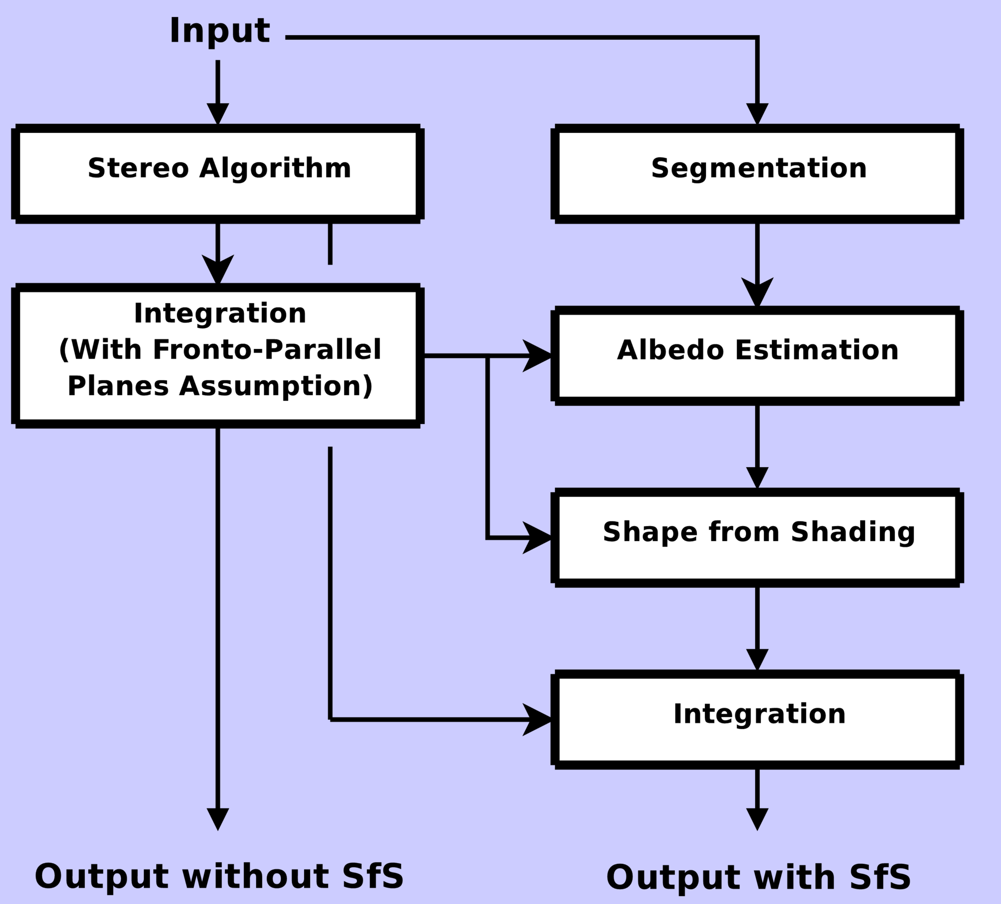
Dept. of Computer Science, THE UNIVERSITY of York, UK.

{thaines,wilson}@cs.york.ac.uk

Abstract

- Binocular Stereo algorithms use correspondence to derive depth information. In practise they will use a smoothing term to compensate for smoothly shaded areas where correspondence fails. Such a term usually assumes {fronto-parallel} piecewise planar surfaces.
- In areas with a known shading model shape-from-shading [SfS] can be used to derive per-pixel orientation information, modelling the surface as being curved rather than planar.
- Our goal is to integrate these two sources of information, depth from stereo and orientation from SfS. The aim is to generate results better than either approach alone.
- We work with disparity from stereopsis and change-in-disparity calculated from SfS. Using Gaussian belief propagation of a Gaussian-Markov random field we find the MAP estimate of the disparity.
- Stereopsis and SfS are performed with dedicated algorithms. A Lambertian model is assumed for SfS, with per-segment albedo estimated using an initial surface estimate derived from stereopsis alone. A single known light source is assumed.
- We compare the results for two scenes with ground truth, in each case running the algorithm using orientation obtained from SfS and orientation set to indicate fronto-parallel planes. An improvement is measured when using SfS information. A qualitative improvement is also observable.

Method



- The *Integration* process is detailed to the right.
- The *Stereo Algorithm* uses dynamic programming to provide a dense disparity map.
- *Segmentation* uses the mean shift algorithm.
- *Albedo Estimation* estimates an albedo for each segment. It uses the irradiance map and smoothed stereo result to calculate per-pixel albedo. The average is used for each segment.
- *SfS* uses a hard cone constraint and is initialised from the smoothed stereo result. It uses the boundary constraint.

Integration with Belief Propagation

- We integrate the disparity and differential of disparity together with belief propagation on a Gaussian-Markov random field.
- The algorithm runs this step twice, first with the differential of disparity set to zero, making it a smoothing operation; and a second time with the SfS calculated disparity differentials.
- Belief propagation iteratively passes messages representing probability distributions between nodes. In our case these represent the distribution on disparity for individual pixels:

$$m_{ts}^{(n)}(x_s) = \alpha \int_{x_t} \psi_{st}(x_s, x_t)\psi_t(x_t, y_t) \prod_{u \in N/s} m_{ut}^{(n-1)}(x_t)dt$$

$m_{ts}^{(n)}(x_s)$ - message from node t to node s at iteration n.

$\psi_{st}(x_s, x_t)$ - compatibility between nodes, parametrised by SfS.

$\psi_t(x_t, y_t)$ - disparity provided by Stereo Algorithm.

after sufficient iterations the final belief may be calculated:

$$b_t^{(n)} = \alpha\psi_t(x_t, y_t) \prod_{u \in N} m_{ut}^{(n-1)}(x_t)$$

- Gaussian distributions may be represented as:

$$\phi[\mathbf{P}\mu, \mathbf{P}] = \alpha \exp\left[-\frac{1}{2}(\mathbf{x}-\mu)^T \mathbf{P}(\mathbf{x}-\mu)\right]$$

the disparity provided by stereopsis may then be represented as:

$$\psi_t(x_t, y_t) = \phi[\mathbf{P}_t\mu_t, \mathbf{P}_t]$$

if the change in disparity between two adjacent pixels is z_{ts} , and P_n is a precision which reflects the accuracy of the surface normal, then a compatibility function can be given by:

$$\psi_{st}(x_s, x_t) = \phi[\mathbf{P_{st}}\mu_{\mathbf{st}}, \mathbf{P_{st}}] = \phi\left[P_n\begin{pmatrix} -z_{ts} \\ z_{ts} \end{pmatrix}, P_n\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\right]$$

- We can define:

$$P_0 = P_t + \sum_{u \in N/s} P_{ut} \qquad P_0\mu_0 = P_t\mu_t + \sum_{u \in N/s} P_{ut}\mu_{ut}$$

and from the above define the messages passed between adjacent nodes:

$$m_{ts}^{(n)}(x_s) = \phi[\mathbf{P}_{ts}\mu_{ts}, \mathbf{P}_{ts}]$$

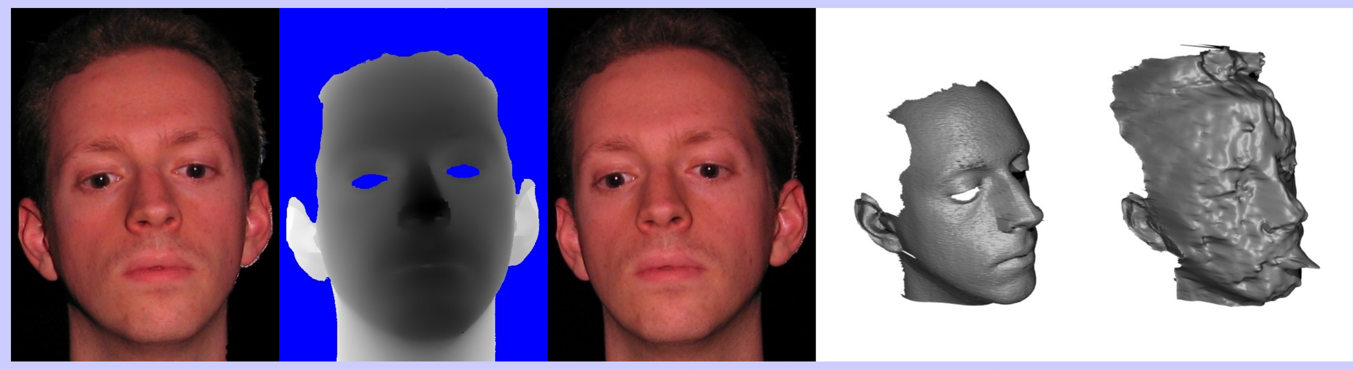
$$P_{ts}\mu_{ts} \leftarrow P_n z_{ts} + P_n(P_n + P_0)^{-1}(P_0\mu_0 - P_n z_{ts})$$

$$P_{ts} \leftarrow P_n - P_n(P_n + P_0)^{-1}P_n$$

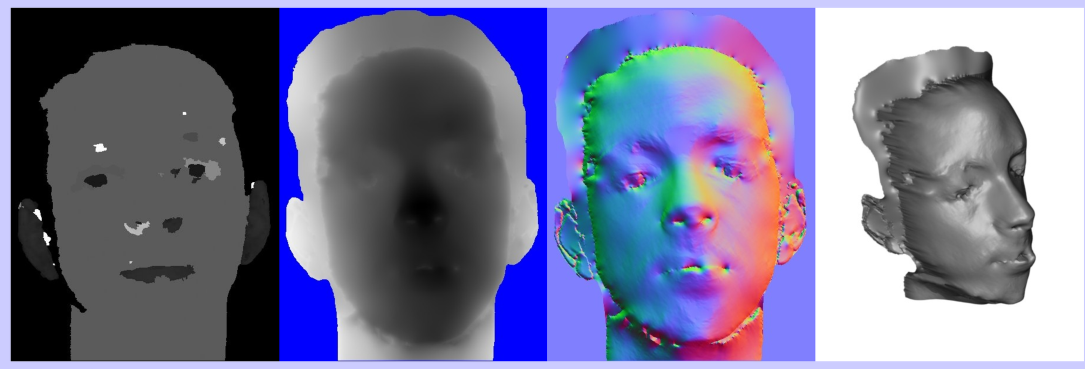
after a sufficient number of iterations the final disparity may be calculated:

$$\mu(t) = (P_t\mu_t + \sum_{u \in N} P_{ut}\mu_{ut})(P_t + \sum_{u \in N} P_{ut})^{-1}$$

Results



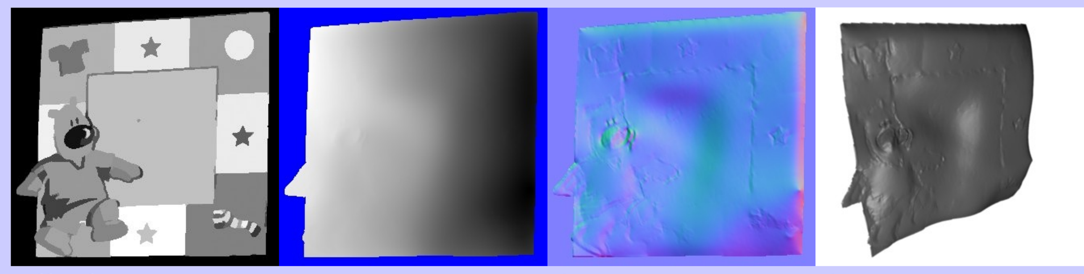
Input, left *Ground Truth* *Input, right* *Ground Truth Render* *Without SfS Render*



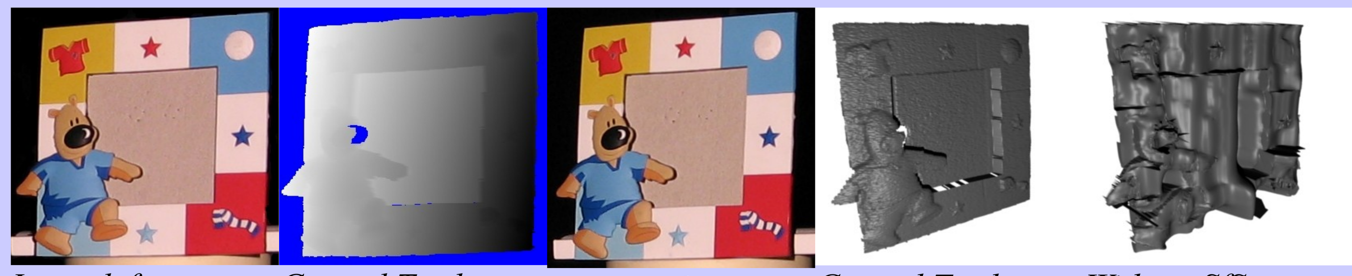
Albedo *Disparity* *Orientation Map* *Output Render*

	Boot Strap	Smoothed Boot Strap	Our Algorithm
Frame	1.62 (13.5%)	1.22 (2.2%)	1.08 (1.0%)
Head	1.84 (13.7%)	1.55 (0.2%)	1.90 (1.9%)
Head without edges	1.73 (10.0%)	1.47 (0.1%)	1.40 (0%)

- The table gives a quantitative analysis. The first number is the average error for inliers; the second is the percentage of outliers. An outlier is defined as a disparity that deviates more than 8 pixels from ground truth.
- For the head image Lambertian reflectance is a bad model at the edges. An improvement is only seen if these areas are ignored, as in the last row of the table.



Albedo *Disparity* *Orientation Map* *Output Render*



Input, left *Ground Truth* *Input, right* *Ground Truth Render* *Without SfS Render*

Conclusions

- We have shown a relatively simple method of combining steropsis and SfS.
- The combined results are shown to be quantitatively better where the SfS works. The output is also perceptually better.
- The quality of SfS, mostly due to albedo estimation, and the ability of the stereopsis to handle smoothly shaded areas appear to be limiting factors. Improved algorithms, probably with greater integration, are required.
- Requiring knowledge of the lighting is a major restriction.
- Orientation sources other then SfS could be considered.