Nanjing University

ICIMCS 2014, Paper ID: 59

Fast Binocular Depth Inference via Bidirectional Motion Based Interpolation

Wenjing Geng, Yang Yang, Ran Ju, Tongwei Ren, Gangshan Wu
State Key Laboratory for Novel Software Technology

Friday July 11, 2014
No. P02

Problem

Acquiring depth sequences specially for stereo videos has seldom been discussed.

Two differences between stereo videos and stereo images

- Adjacent frames for a natural video are highly correlated.
- Frames in one shot have implicit continuity.

How to employ the content redundancy and consistency between frames?
How to preserve the continuity in depth sequences?

Related work

[Hung et al., IJCV 2013]: time complexity is high.
[Valgaerts et al., ECCV 2010]: four frame configuration.

Solution

Basic idea: Try to leverage the inter-frame redundancy to reduce computational time and preserve consistency

Framework

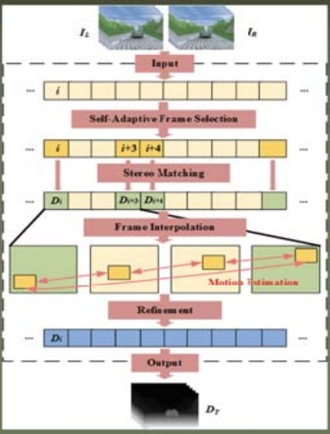

Figure 1: The framework of the proposed approach

$$decision(i) = \begin{cases} 1, & \|P_i^b - P_i^a\| > \tau \\ 0, & \text{otherwise} \end{cases}$$

$$D_K(x+\alpha u, y+\alpha v) = \alpha D_m(x,y) + (1-\alpha) D_n(x+u, y+v)$$
$$\alpha = \frac{k-m}{n-m}$$

$$assignment(i) = \begin{cases} 0 & d_i = 0 \text{ and } \sum_{j \in N(i)} I(d_j) < \epsilon \\ 1 & \text{otherwise} \end{cases}$$
$$I(x) = \begin{cases} 1 & x = 0 \\ 0 & x > 0 \end{cases}$$
$$D_k(x,y) = \beta D_m(x,y) + (1-\beta) D_n(x,y)$$
$$\beta = \frac{k-m}{n-m}$$

Experiment

Comparison

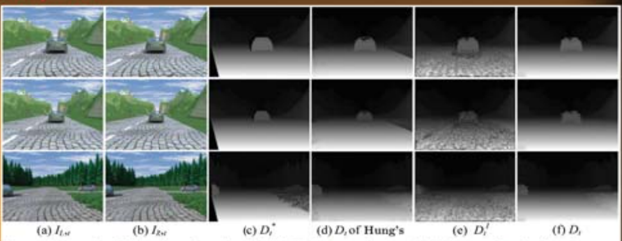

Figure 2: Example of disparity estimation. (a)-(b) Color stereo images. (c) The ground truth disparity map. (d) The estimated disparity map from Hung's. (e)-(f) Our initial and final inference results.

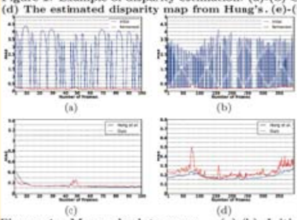

Figure 4: Mean absolute error. (a)-(b) Initial and refinement results of our method. (c)-(d) Comparison with the curve from Hung's.

Table 1: Comparison of running time (in minutes)

	Total	Per frame
Hung et al. (single thread)	2772	28
Hung et al. (multi thread)	138	1.4
Ours (single thread)	51.606	0.516

Table 2: Running time (in minutes) of our approach

	OFBSM	MBI	Ref.	Total	Per Frame
Scene1	44.437	7.126	0.043	51.606	0.516
Scene2	294.025	16.675	0.099	310.799	0.785
Scene3	82.434	12.170	0.061	94.665	0.845

(note: OFBSM is optical flow based stereo matching, MBI is motion based interpolation, and Ref. is short for refinement)

Figure 3: Our inference disparity map from Karlsruhe dataset. (a) Left frame from stereo video. (b) Inference results of the proposed approach.

Conclusion

Contribution

- Propose a fast binocular depth inference framework for applications of time requirement.
- A self-adaptive strategy for selecting frames is employed to control the interpolation error propagation within a proper range.
- An interpolation method based on bidirectional motion estimation is presented to guarantee the consistency and accuracy of the adjacent frames between selected frames

Future work

Improving the framework and applying our method to further multimedia applications, such as 3D reconstruction and scene understanding.

Contact to: jenngeng@gmail.com