Summary

Interest: (R1) Clearly, this should allow very differents image modalities to be matched, as long as there is scene motion to be measured.

(R3) The paper addresses an interesting topic of cross modality image correspondence estimation for 3D reconstruction and 3D reconstruction for coaxial cameras.

Positives:

Innovative/Novel: (R1) The idea of using a coaxial setup for stereo combining the optical flow for matching is innovative, in my opinion. The derivation which links disparity, stereo geometry and optical flow is interesting and might find other use for non coaxial cameras.

(R2) The idea of performing disparity estimation from optical flow input rather than image correspondences seems original.

(R2) Technically, the main novelty is the introduction of the constraints in equations (1) and (6) - though I did not check the correctness as derivations have not been provided.

(R3) the paper introduces a novel approach for estimating correspondences not directly leveraging visual correspondence but image sequence motion.

(R3) The paper presents a rather new stereo matching approach, where the optical flows in both images are matched rather than the intensity.

Technical Interest: (R3) the method is technically sound, (R1) the paper propose interesting ideas to tackle it (especially matching the optical flow). (R3) On one hand it has a clear contribution and a sound theory behind it

Well written (R1) The paper is clearly written an reproduceable.

Weaknesses:

(R1) Experimental results, synthetic scene not well described, RMS error has limited usefulness. Wants to see the IR/RGB image pair to appreciate difference. Wants to see the optical flow field. Didn't detect it, but the traditional stereo rig has problems with large planar surfaces parallel to the camera plane (e.g. there is not variation in the optical flow), the surface needs "mountains".

(R2) Missing derivations. Could be provided in supplementary material. Convergence, Limited number of scenes considered, lack of comparison against other techniques Application/scope (this is in the paper).

(R3) More real data. Images of scenes,

(R4) Initialization, Issues with frontal parallel planes, Missing numerical method used to solve the problem. Needs comparison with state-of-the-art multimodal stereo matching method. RMS error need more explaination.

Some misunderstandings relative to intialization, baseline, comparisons

Issues:

Initialization.

Better comparison with State

Suggestions: (R1) Provide the image sequences