INRIA

PRIMA Team
Laboratoire Informatique de Grenoble
UMR 5217

Mail: {sergi.pujades-rocamora,frederic.devernay}@inria.fr
Inria Grenoble - Rhône Alpes
655 Av de l'Europe 38330 Montbonnot-Saint-Martin

Stereoscopic Focus Mismatch Monitoring

Sergi Pujades Rocamora - Frédéric Devernay

Abstract

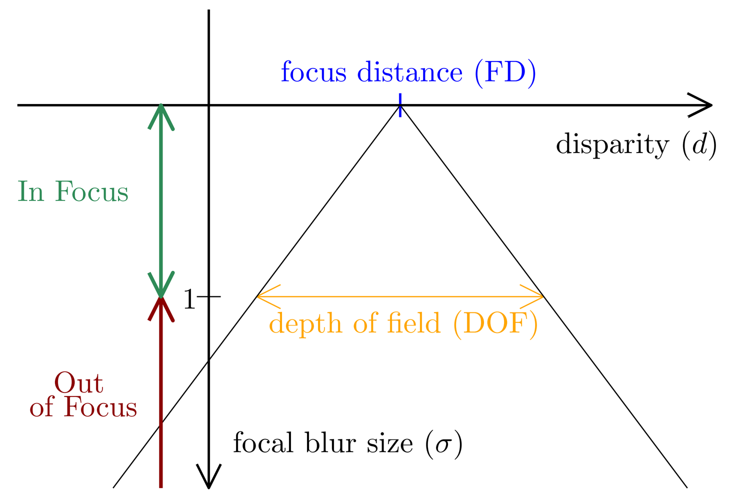
We detect focus mismatch between views of a stereoscopic pair. First, we compute a dense disparity map. Then, we use a measure to compare focus in both images. Finally we use robust statistics to find which images zones have different focus. We show the results on the original images.

Introduction

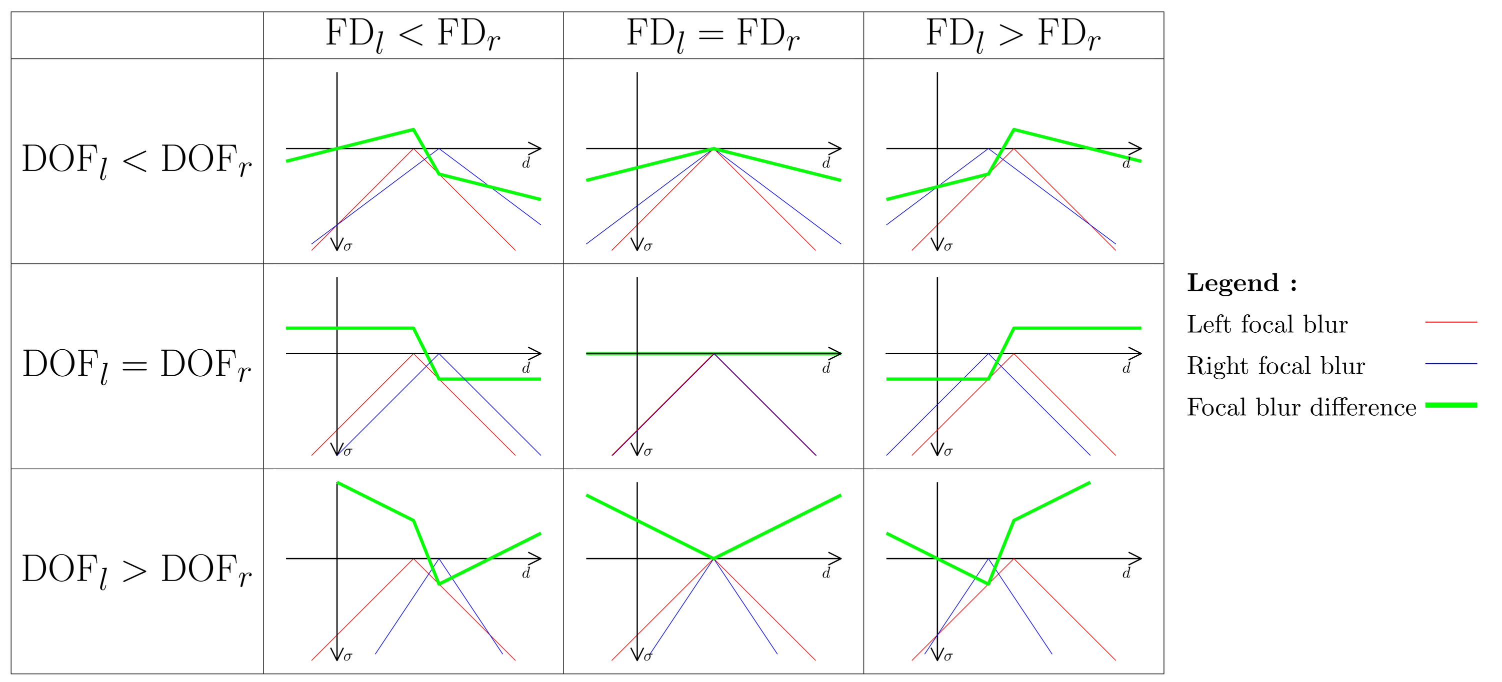
Live-action stereoscopic content production requires a stereo rig with two cameras precisely matched and aligned. While most deviations from this perfect setup can be corrected either live or in post-production, a difference in the focus distance or focus range between the two cameras will lead to unrecoverable degradations of the stereoscopic footage.

Measuring the focus at a point in an image is an ill-posed problem, since it may be the in-focus image of a non-textured point in the scene. We consider that the captured image is the result of the application of the focus blur on an image with infinite depth of field.

Assuming parallel (or near-parallel) cameras the focus value should be dependent on depth, and thus on disparity. It can be shown that the focal blur size is linear with the stereo disparity, as seen in the figure below.



The difference of focal blur between two images, as a function of disparity, has one of the following nine different shapes.



Despite the fact that we cannot measure the focal blur directly, but only image blur, the sign of image blur difference has the same sign as the focal blur difference. We can thus use this function to measure focus mismatch.

Algorithm Outline

1. Compute dense disparity map.
2. Measure focus mismatch.
3. Model the focus mismatch measures.
4. Draw mismatched areas on the original images.

Disparity map computation

We first compute a dense disparity map. Since images may differ in focus, we use a real-time multi-scale method [3] which finds good disparity values even between focused and blurred textures. We also compute semi-occluded areas by left-right consistency check, and ignore them in the following computation. Let $d(\mathbf{i})$ be the disparity of pixel $\mathbf{i} = (x, y)$.

Focus Mismatch Measurement

In order to compare the focus of corresponding points from the left and right images we use the SML operator (sum of modified Laplacian [1, 2]) which was primarily designed for depth-from-focus applications,

$$\mathrm{SML}(\mathbf{i}) = \sum_{r=x-N}^{x+N} \sum_{s=y-N}^{y+N} \nabla^2_M(r,s), \text{ for } \nabla^2_M(r,s) \geq T_1, \text{ where } \nabla^2_M I = \left|\frac{\partial^2 I}{\partial x^2}\right| + \left|\frac{\partial^2 I}{\partial y^2}\right|.$$

For a pixel $\mathbf{i}$ in the left image, $\mathrm{SML}_l(\mathbf{i})$ is the SML operator computed at this pixel, and $\mathrm{SML}_r(\mathbf{i})$ is computed at the corresponding pixel in the right image.

Let $M(\mathbf{i})$ be the sign of the difference of SML between two pixels:

$$M(\mathbf{i}) = \mathrm{sign}(\mathrm{SML}_l(\mathbf{i}) - \mathrm{SML}_r(\mathbf{i})).$$

It is positive if left is more focused than right, and negative if right is more focused than left.

Let $w(\mathbf{i})$ be the max of $\mathrm{SML}_l(\mathbf{i})$ and $\mathrm{SML}_r(\mathbf{i})$:

$$w(\mathbf{i}) = \max(|\mathrm{SML}_l(\mathbf{i})|, |\mathrm{SML}_r(\mathbf{i})|).$$

It is high if one image is textured and in focus, which tells us where reliable information is.

For each disparity value of the scene, $M(d)$ is the mean of those differences at disparity d weighted by $w(\mathbf{i})$, and $w(d)$ is the corresponding sum of weights. This gives us an estimate of which image is more focused at disparity d.

However this estimate may be very noisy :

- The disparity information may be inaccurate, which leads $M(\mathbf{i})$ to be wrong.
- Focal blur bleeds over depth discontinuities : background objects may be wrongly measured as blurred.
- The number of measures for each disparity depends highly on scene content and information may even be missing for some disparity values.

In order to tackle these problems we perform a model estimation to obtain robust information.

Focus difference model estimation

Given the measures $M(d)$ and $w(d)$ we fit a model $C(d)$ minimising the energy

$$E = \sum_d w(d) E_{\text{data}}(d) + \lambda E_{\text{smooth}}(d), \text{ with}$$

$$E_{\text{data}}(d) = M(d) - C(d) \text{ and } E_{\text{smooth}}(d) = |C(d-1) - C(d)|.$$

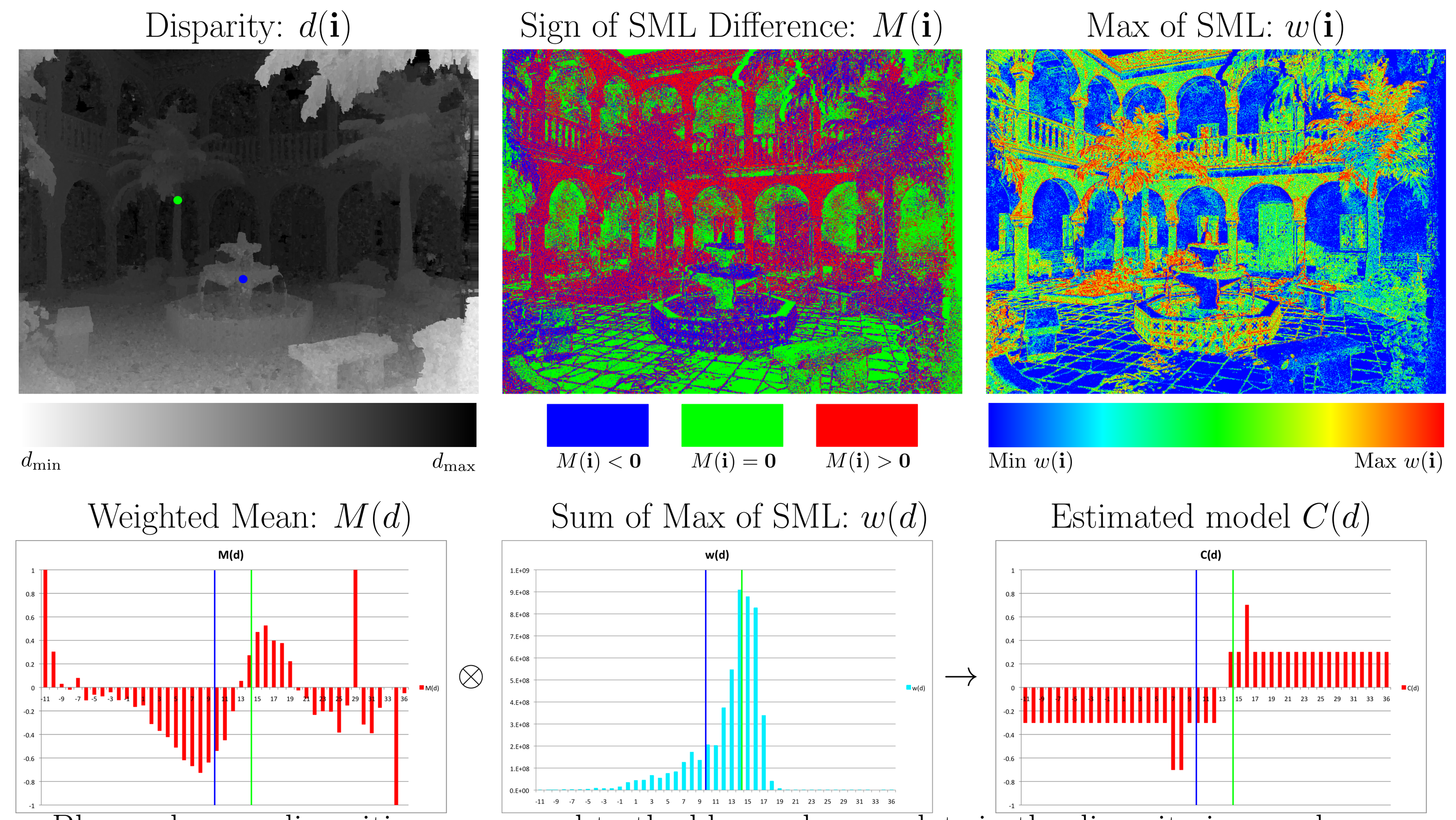
$C(d)$ may take 5 possible values:

- Same Focus ($= 0$)
- Left slightly more in focus (≈ 0.3)
- Left more in focus (≈ 0.7)
- Right slightly more in focus (≈ -0.3)
- Right more in focus (≈ -0.7)

The fact that the sign of the focus difference must correspond to one of the nine shapes described before brings one more constraint to the model: the number of sign changes is limited. The minimization can be solved in linear time ($O(d_{\max} - d_{\min})$) using dynamic programming. The result is a curve telling for each disparity if left is more focused than right, both focus are the same, or right is more focused than left.

Results

The results below use a pair of ray-traced images with the same depth of field and different focus distances: left image is focused on the green dot, right image is focused on the blue dot.



Blue and green disparities correspond to the blue and green dots in the disparity image above.

To graphically view the result we draw zebra strokes on each original image, marking the areas where this image is less focused than the other.

Left red zebra area (e.g. the fountain) is less focused than in the right image.

Right red zebra area (e.g. the background) is less focused than in the left image.

A closer look at a detail of the original images visually confirms our results.

Left detail — Right detail

References

[1] Wei Huang and Zhongliang Jing. Evaluation of focus measures in multi-focus image fusion. Pattern Recognition Letters, 28(4):493 – 500, 2007.

[2] S.K. Nayar and Y. Nakagawa. Shape from focus. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 16(8):824 –831, aug 1994.

[3] M. Sizintsev, S. Kuthirummal, H. Sawhney, A. Chaudhry, S. Samarasekera, and R. Kumar. Gpu accellerated realtime stereo for augmented reality. In Proceedings Intl. Symp. 3D Data Processing, Visualization and Transmission (3DPVT), 2010.

This work was done within the 3DLive project supported by the French Ministry of Industry
http://3dlive-project.com/