Motion Parallax in Stereo 3D: Model and Applications

Supplemental Materials

Petr Kellnhofer[1] Piotr Didyk[1,2] Tobias Ritschel[1,2,3] Belen Masia[1,4] Karol Myszkowski[1] Hans-Peter Seidel[1]

[1]MPI Informatik [2]Saarland University, MMCI [3]University College London [4]Universidad de Zaragoza

Abstract

This is a supplemental text for the SIGGRAPH Asia 2016 paper Motion Parallax in Stereo 3D: Model and Applications. It provides additional information that did not fit in the paper.

Keywords: stereoscopic 3D, gaze tracking, gaze-contingent display, eye tracking, retargeting, remapping

Concepts: •**Computing methodologies** → **Perception;** *Image processing;*

1 Numerical example

Here we provide a numerical example of our model application for two 3D world points $\mathbf{x}_A$ and $\mathbf{x}_B$.

1.1 Problem statement

The point $\mathbf{x}_A$ is located at 3D coordinates $\mathcal{P}(\mathbf{x}_A) = [0, 0, 2.0]$ in meters and point $\mathbf{x}_B$ is located at $\mathcal{P}(\mathbf{x}_B) = [0.01, 0, 2.1]\,\mathrm{m}$ (see Fig. 1). We assume that the world is static and the camera at coordinates $C = [0, 0, 0]$ m is facing in the z-axis direction and moves laterally in the x-axis direction with a velocity $h = 30\,\mathrm{cm/s}$. The scene is displayed on a screen at distance $l = 0.6\,\mathrm{m}$ from an observer. The two points are projected on the screen with a separation of $r = 3$ pixels where the pixel size is $p = 0.257\,\mathrm{mm}$. The distance between the eyes of the observer is $i = 6.4\,\mathrm{cm}$. A linear compression with a factor $c = 0.2$ was applied to the content in order to fit it to the display depth budget.

Our goal is to compute the disparity between points $\mathbf{x}_A$ and $\mathbf{x}_B$ that in combination with motion parallax produces perceived depth magnitude equal to the depth from disparity in a single static frame. We achieve this by solving Eq. 15 in the paper.

1.2 Model specification

First, we express the problem in terms of the model variables. We have to find relative depth from motion parallax m, binocular disparity d, spatial separation s and retinal velocity V.

SA '16 Technical Papers, December 05 - 08, 2016, Macao
ISBN: 978-1-4503-4514-9/16/12
DOI: http://dx.doi.org/10.1145/2980179.2980230

Relative depth from motion parallax $m = \Delta\mathcal{D}_m$ First, we express distances from the observer to both points:

$$f_A = ||P_A - C|| = ||[0, 0, 2.0] - [0, 0, 0]|| = 2$$
$$f_B = ||P_B - C|| = ||[0.01, 0, 2.1] - [0, 0, 0]|| \approx 2.1.$$

Next, we can directly use our definition of $m = \Delta\mathcal{D}_m$:

$$\Delta\mathcal{D}_m = \frac{\Delta f}{f} = \frac{|f_A - f_B|}{\max(f_A, f_B)} = \frac{|2.0 - 2.1|}{\max(2.0, 2.1)} = \frac{0.1}{2.1} \approx 0.048.$$

Binocular disparity $d = \Delta\mathcal{D}_d$ First, we express the vergence of the points in the uncompressed 3D world:

$$\begin{aligned}\mathcal{D}_d(A) &= 2\arctan\frac{i}{2f_A}\\ &= 2\arctan\frac{0.064}{2\cdot 2.0} \approx 0.0320\,\mathrm{rad} \approx 110.0\,\mathrm{arcmin}\\ \mathcal{D}_d(B) &= 2\arctan\frac{i}{2f_B}\\ &= 2\arctan\frac{0.064}{2\cdot 2.1} \approx 0.0305\,\mathrm{rad} \approx 104.8\,\mathrm{arcmin}.\end{aligned}$$

Next, we compute the disparity $d = \Delta\mathcal{D}_d$ while accounting for the compression factor c as:

$$\Delta\mathcal{D}_d = |\mathcal{D}_d(\mathbf{x}_A) - \mathcal{D}_d(\mathbf{x}_B)| \cdot c = |110.0 - 104.8| \cdot 0.2 = 1.04\,\mathrm{arcmin}.$$

Spatial separation $s = \Delta\mathcal{D}_s$ We use information about the pixel distance and convert it to an angular measure:

$$\begin{aligned}\Delta\mathcal{D}_s &= 2\arctan\frac{r\cdot p}{2l}\\ &= 2\arctan\frac{3\cdot 0.257\cdot 10^{-3}}{2\cdot 0.6} \approx 0.0013\,\mathrm{rad} \approx 4.42\,\mathrm{arcmin}.\end{aligned}$$

Retinal velocity $V = \Delta\mathcal{D}_v$ First, we find angular velocities of each of the two points. Typically, we would use the difference of projections at two different times. For the purpose of this example and the sake of simplicity, we derive both velocities in the 3D world and then apply the field of view compression of the display. In the 3D world, the angular velocities are based on the distance and the observer's motion:

$$\begin{aligned}\nu_A &\approx \arctan\frac{h}{f_A}\\ &= \arctan\frac{0.3}{2.0} \approx 0.1489\,\mathrm{rad/s} \approx 511.8\,\mathrm{arcmin/s}\\ \nu_B &\approx \arctan\frac{h}{f_B}\\ &= \arctan\frac{0.3}{2.1} \approx 0.1419\,\mathrm{rad/s} \approx 487.8\,\mathrm{arcmin/s}.\end{aligned}$$

The scaling from the 3D world to the screen coordinates is derived from the ratio of the angular separation of both points in the screen

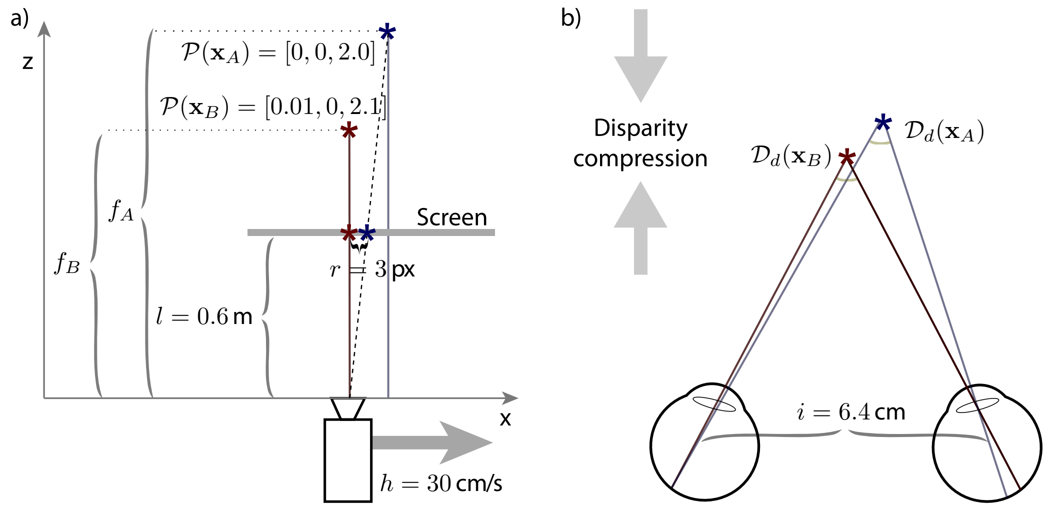

Figure 1: *A two-point setup used in the numerical example. a) The 3D world as viewed by the (depth) camera illustrating physical layout and motion which jointly determine the motion parallax in the scene. b) The stereoscopic presentation on the screen after a linear disparity compression as viewed by the observer.*

space $\Delta\mathcal{D}_s$ and the same separation in the 3D world $\Delta\mathcal{D}_s^{\text{3D}}$. Given our scene configuration:

$$\begin{aligned}\Delta\mathcal{D}_s^{\text{3D}} &\approx 2\arctan\frac{|\mathcal{P}(A)^x - \mathcal{P}(B)^x|}{2|\mathcal{P}(B)^z - C^z|}\\ &= 2\arctan\frac{|0-0.01|}{2|2.1-0|} \approx 0.0048\,\text{rad} \approx 16.37\,\text{arcmin}.\end{aligned}$$

Finally, the retinal velocity $V = \Delta\mathcal{D}_v$ between $\mathbf{x}_A$ and $\mathbf{x}_B$ is:

$$\begin{aligned}\Delta\mathcal{D}_v &= |\nu_A - \nu_B| \cdot \frac{\Delta\mathcal{D}_s}{\Delta\mathcal{D}_s^{\text{3D}}}\\ &= |511.8 - 487.8| \cdot \frac{4.42}{16.37} \approx 6.48\,\text{arcmin/s}.\end{aligned}$$

1.3 Parallax map

The first step towards the disparity scaling problem is evaluation of the parallax map $M(\mathbf{x})$ according to Eq. 14. This will tell us whether the necessary conditions for existence and visibility of motion parallax are fulfilled. According to the algorithm, we should first specify the point $\mathbf{x}'$ using Eq. 13. However, in this example we have only two points, and therefore we choose $\mathbf{x} = \mathbf{x}_A$ and $\mathbf{x}' = \mathbf{x}_B$.

Next, we evaluate the algorithm step by step starting from Eq. 8 for transformation matrix similarity metric $\gamma(\mathbf{x}_A, \mathbf{x}_B)$:

$$\gamma(\mathbf{x}_A, \mathbf{x}_B) = \frac{1}{|\mathcal{N}|}\sum_{\mathbf{x}_i \in \mathcal{N}} \|\mathcal{M}(\mathbf{x}_A)\,\mathcal{P}(\mathbf{x}_i) - \mathcal{M}(\mathbf{x}_B)\,\mathcal{P}(\mathbf{x}_i)\|_2.$$

In our case, the transformation matrix map $\mathcal{M}(\mathbf{x})$ is constant, as we have defined the scene to be static and therefore rigid. The only motion originates from the camera motion, which can be described by a translation:

$$\mathcal{M}(\mathbf{x}) = \mathcal{M} = \begin{pmatrix}1 & 0 & 0 & h\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix} = \begin{pmatrix}1 & 0 & 0 & 0.3\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}.$$

For our simplified neighborhood $\mathcal{N}$ with only two points we get a trivial expression:

$$\begin{aligned}\gamma(\mathbf{x}_A, \mathbf{x}_B) &= \frac{1}{2}\|\mathcal{M}(\mathbf{x}_A)\,\mathcal{P}(\mathbf{x}_A) - \mathcal{M}(\mathbf{x}_B)\,\mathcal{P}(\mathbf{x}_A)\|_2\\ &\quad + \|\mathcal{M}(\mathbf{x}_A)\,\mathcal{P}(\mathbf{x}_B) - \mathcal{M}(\mathbf{x}_B)\,\mathcal{P}(\mathbf{x}_B)\|_2\\ &= \frac{1}{2}\|\mathcal{M}\cdot[0,0,2.0] - \mathcal{M}\cdot[0,0,2.0]\|_2\\ &\quad + \|\mathcal{M}\cdot[0.01,0,2.1] - \mathcal{M}\cdot[0.01,0,2.1]\|_2\\ &= 0.\end{aligned}$$

The rigidity measure $\Gamma(\mathbf{x}_A, \mathbf{x}_B)$ is then equal to (Eq. 9):

$$\begin{aligned}\Gamma(\mathbf{x}_A, \mathbf{x}_B) &= \begin{cases}1 & \text{when } \gamma(\mathbf{x}_A, \mathbf{x}_B) < \epsilon,\\ 0 & \text{otherwise,}\end{cases}\\ &= \begin{cases}1 & \text{when } 0 < \epsilon,\\ 0 & \text{otherwise,}\end{cases}\\ &= 1.\end{aligned}$$

This confirms that the transformation is rigid.

According the Eq. 10, the visibility of motion parallax $\zeta(\mathbf{x}_A, \mathbf{x}_B)$ is equal to:

$$\begin{aligned}\zeta(\mathbf{x}_A, \mathbf{x}_B) &= \Psi(\Gamma(\mathbf{x}_A, \mathbf{x}_B)\cdot\Delta\mathcal{D}_v, \Delta\mathcal{D}_s)\\ &= \Psi_1\left(\frac{1.25}{\Delta\mathcal{D}_s}\Gamma(\mathbf{x}_A, \mathbf{x}_B)\cdot\Delta\mathcal{D}_v\right)\\ &= \begin{cases}1 & \text{when } \frac{1.25}{\Delta\mathcal{D}_s}\Gamma(\mathbf{x}_A, \mathbf{x}_B)\cdot\Delta\mathcal{D}_v \geq v_c,\\ 0 & \text{otherwise}\end{cases}\\ &= \begin{cases}1 & \text{when } \frac{1.25\cdot 60}{4.42}\cdot 1\cdot 6.48 \geq v_c,\\ 0 & \text{otherwise}\end{cases}\\ &= \begin{cases}1 & \text{when } 109.95 \geq 2.36,\\ 0 & \text{otherwise}\end{cases}\\ &= 1,\end{aligned}$$

where $\Psi(V, s)$ is the retinal velocity coefficient function fitted as part of our model in Eq. 7. This shows that there is a visible motion between the two points according to our model.

Finally, the expression for parallax map $M(\mathbf{x})$ according the Eq. 14 can be evaluated:

$$\begin{aligned} M(\mathbf{x}_A) = M &= \Gamma(\Delta\mathbf{x}_A) \cdot \zeta(\mathbf{x}_A, \mathbf{x}_B) \\ &= \Gamma(\mathbf{x}_A, \mathbf{x}_B) \cdot \zeta(\mathbf{x}_A, \mathbf{x}_B) \\ &= 1 \cdot 1 \\ &= 1, \end{aligned}$$

which puts both conditions together and predicts the existence of perceivable motion parallax.

1.4 Disparity scaling

We have resolved that there is a visible motion parallax between points $\mathbf{x}_A$ and $\mathbf{x}_B$, and we can decide about its contribution to the perceived depth when combined with the present disparity. We do this by solving Eq. 15, which encapsulates our perceptual model:

$$\forall_{\mathbf{x}}\, \Delta\mathcal{D}_d(\mathbf{x}_A) = \Phi\left(\Delta\mathcal{D}_m(\mathbf{x}_A), \Delta\mathcal{D}'_d(\mathbf{x}_A), M(\mathbf{x}_A)\right).$$

We optimize only for disparity between our two points and given the symmetry of this relation, we can simplify this to:

$$\Delta\mathcal{D}_d = \Phi\left(\Delta\mathcal{D}_m, \Delta\mathcal{D}'_d, M\right),$$

where $\Phi(m, d, \Psi(.))$ is our full model from Eq. 4:

$$\Delta\mathcal{D}_d(\mathbf{x}) = \Delta\mathcal{D}'_d(\mathbf{x}) + M \cdot \left(\Phi_{\hat{V}}(\Delta\mathcal{D}_m, \Delta\mathcal{D}'_d) - \Delta\mathcal{D}'_d\right),$$

and $\Phi_{\hat{V}}(m, d)$ is the supra-threshold velocity model from Eq. 2. After expansion we obtain a monotonic non-linear equation with single unknown $\Delta\mathcal{D}'_d$:

$$\begin{aligned} \Delta\mathcal{D}_d &= \Delta\mathcal{D}'_d \\ &\quad + M \cdot \left(\left(c_1 \cdot \lambda(\Delta\mathcal{D}'_d) + c_2\right) \cdot \Delta\mathcal{D}_m + c_3 \cdot \Delta\mathcal{D}'_d - \Delta\mathcal{D}'_d\right) \\ \Delta\mathcal{D}_d &= \Delta\mathcal{D}'_d \\ &\quad + M \cdot \left(\left(c_1 \cdot \left(\frac{2}{1 + e^{-\Delta\mathcal{D}'_d \beta}} - 1\right) + c_2\right) \cdot \Delta\mathcal{D}_m \right. \\ &\qquad \left. + c_3 \cdot \Delta\mathcal{D}'_d - \Delta\mathcal{D}'_d\right) \\ 4.42 &= \Delta\mathcal{D}'_d \\ &\quad + 1 \cdot \left(\left(-50.88 \cdot \left(\frac{2}{1 + e^{-\Delta\mathcal{D}'_d \cdot 0.25}} - 1\right) + 68.56\right) \cdot 0.048 \right. \\ &\qquad \left. + 1.006 \cdot \Delta\mathcal{D}'_d - \Delta\mathcal{D}'_d\right), \end{aligned}$$

which we solve numerically in a fixed number of steps by bisection to get $\Delta\mathcal{D}'_d \approx 1.60$ arcmin.

Conclusion By evaluating Eq. 14 for M we have verified that the transformation of both points with respect to the camera is rigid and that their mutual retinal velocity from the observer's perspective is sufficient for a perceivable motion parallax. The result of the optimization in Eq. 15 told us that in order to represent the same perceived depth as in a static frame with binocular disparity $\Delta\mathcal{D}_d = 4.42$ arcmin we only need a disparity magnitude $\Delta\mathcal{D}'_d = 1.60$ arcmin if a depth cue from motion parallax for relative depth $\Delta\mathcal{D}_m = 0.048$ is also available.

2 Experiment statistics

Table 1 shows statistics for individual measurement points from our model construction experiment in Sec. 3 of the main paper. We provide a complete list of measured values in a separate spreadsheet supplement.

$m = \frac{\Delta f}{f}$ [-]	d [arcmin]	$V = \frac{d\theta}{dt}$ [arcmin/s]	$\nu = \frac{d\alpha}{dt}$ [arcmin/s]	m_e [arcmin]	f [cm]	Δf [cm]	Reference disparity [arcmin] Mean	SEM	SD	Sample count
0.000	0	0.001	500	0.00	60.00	0.00	0.683	0.144	1.059	54
0.000	5	0.001	500	0.00	60.00	0.00	6.130	0.303	2.223	54
0.000	10	0.001	500	0.00	60.00	0.00	9.980	0.431	3.168	54
0.000	15	0.001	500	0.00	60.00	0.00	15.431	0.453	3.330	54
0.000	20	0.001	500	0.00	60.00	0.00	21.091	0.756	5.557	54
0.027	0	13.138	500	10.00	60.82	1.64	2.665	0.369	2.709	54
0.027	5	13.138	500	10.00	60.82	1.64	6.311	0.297	2.181	54
0.027	10	13.138	500	10.00	60.82	1.64	10.315	0.311	2.288	54
0.027	15	13.138	500	10.00	60.82	1.64	15.420	0.684	5.075	55
0.027	20	13.138	500	10.00	60.82	1.64	20.120	0.810	5.951	54
0.053	0	25.307	500	20.03	61.64	3.28	4.039	0.475	3.488	54
0.053	5	25.307	500	20.03	61.64	3.28	7.148	0.367	2.694	54
0.053	10	25.307	500	20.03	61.64	3.28	11.437	0.514	3.774	54
0.053	15	25.307	500	20.03	61.64	3.28	15.633	0.657	4.828	54
0.053	20	25.307	500	20.03	61.64	3.28	21.300	0.961	7.065	54
0.079	0	36.637	500	30.10	62.47	4.93	6.978	0.899	6.604	54
0.079	5	36.637	500	30.10	62.47	4.93	9.972	0.827	6.074	54
0.079	10	36.637	500	30.10	62.47	4.93	12.561	0.711	5.225	54
0.079	15	36.637	500	30.10	62.47	4.93	17.337	0.965	7.089	54
0.079	20	36.637	500	30.10	62.47	4.93	23.376	1.167	8.579	54
0.104	0	47.226	500	40.24	63.29	6.58	7.511	0.891	6.546	54
0.104	5	47.226	500	40.24	63.29	6.58	10.107	0.839	6.167	54
0.104	10	47.226	500	40.24	63.29	6.58	13.044	0.652	4.791	54
0.104	15	47.226	500	40.24	63.29	6.58	18.064	1.052	7.802	55
0.104	20	47.226	500	40.24	63.29	6.58	24.380	1.438	10.569	54
0.027	0	0.000	0	10.00	60.82	1.64	0.400	0.169	1.016	36
0.027	0	0.263	10	10.00	60.82	1.64	0.631	0.249	1.495	36
0.027	0	1.314	50	10.00	60.82	1.64	0.633	0.178	1.069	36
0.027	0	2.628	100	10.00	60.82	1.64	1.111	0.390	2.342	36
0.027	0	7.883	300	10.00	60.82	1.64	2.486	0.693	4.156	36
0.079	10	0.000	0	30.10	62.47	4.93	9.885	0.330	2.425	54
0.079	10	0.733	10	30.10	62.47	4.93	10.574	0.402	2.958	54
0.079	10	3.664	50	30.10	62.47	4.93	11.833	0.547	4.017	54
0.079	10	7.327	100	30.10	62.47	4.93	12.144	0.585	4.302	54
0.079	10	21.982	300	30.10	62.47	4.93	12.474	0.815	5.988	54
0.104	5	0.000	0	40.24	63.29	6.58	5.411	0.381	2.316	37
0.104	5	0.945	10	40.24	63.29	6.58	8.242	0.563	3.376	36
0.104	5	4.723	50	40.24	63.29	6.58	8.872	0.874	5.242	36
0.104	5	9.445	100	40.24	63.29	6.58	9.816	1.218	7.411	37
0.104	5	28.336	300	40.24	63.29	6.58	9.914	1.108	6.740	37

Table 1: *Statistics of measured equivalent reference disparity for individual measurement points. Note that spatial separation $s = 1.25$ deg was fixed for all stimuli. The mean fitting error is 0.965 arcmin when means for individual measurement points are compared to our model. The mean fitting error for all individual measured values is 3.468 arcmin.*