DEPTH FROM GAZE

TZU-SHENG KUO, KUANG-TSU SHIH, SHENG-LUNG CHUNG, and HOMER CHEN
NATIONAL TAIWAN UNIVERSITY

Introduction

- We exploit the gaze information acquired by an eye tracker for depth estimation.
- The proposed method can be used to construct a sparse depth map of a visual world.

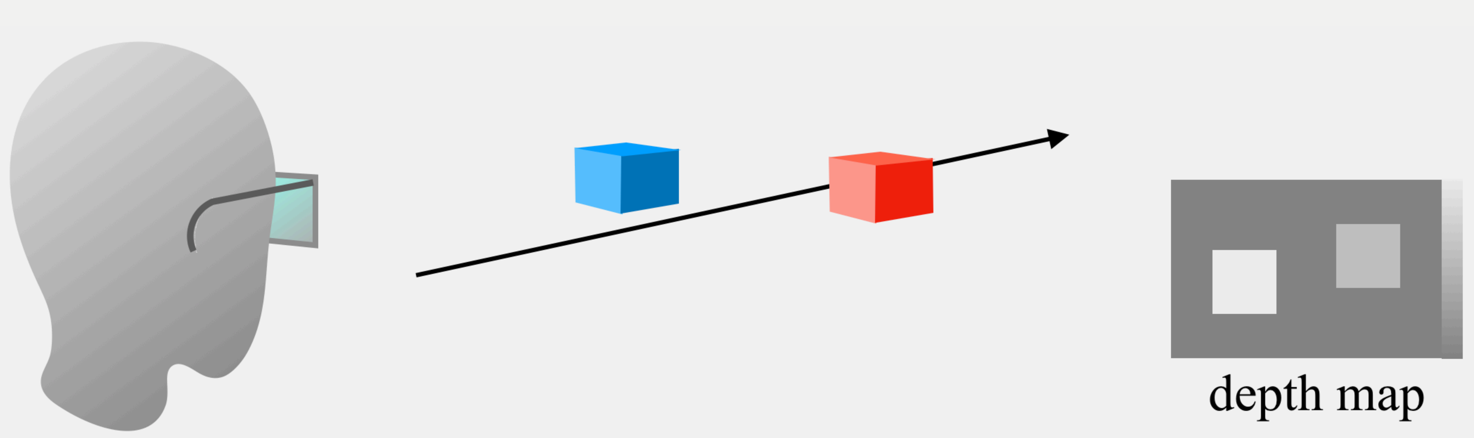


Fig. 1. An application scenario.

Motivation

- We believe that it is possible to address the depth estimation problem using gaze information because human eyeballs rotate when gazing targets at different depths.

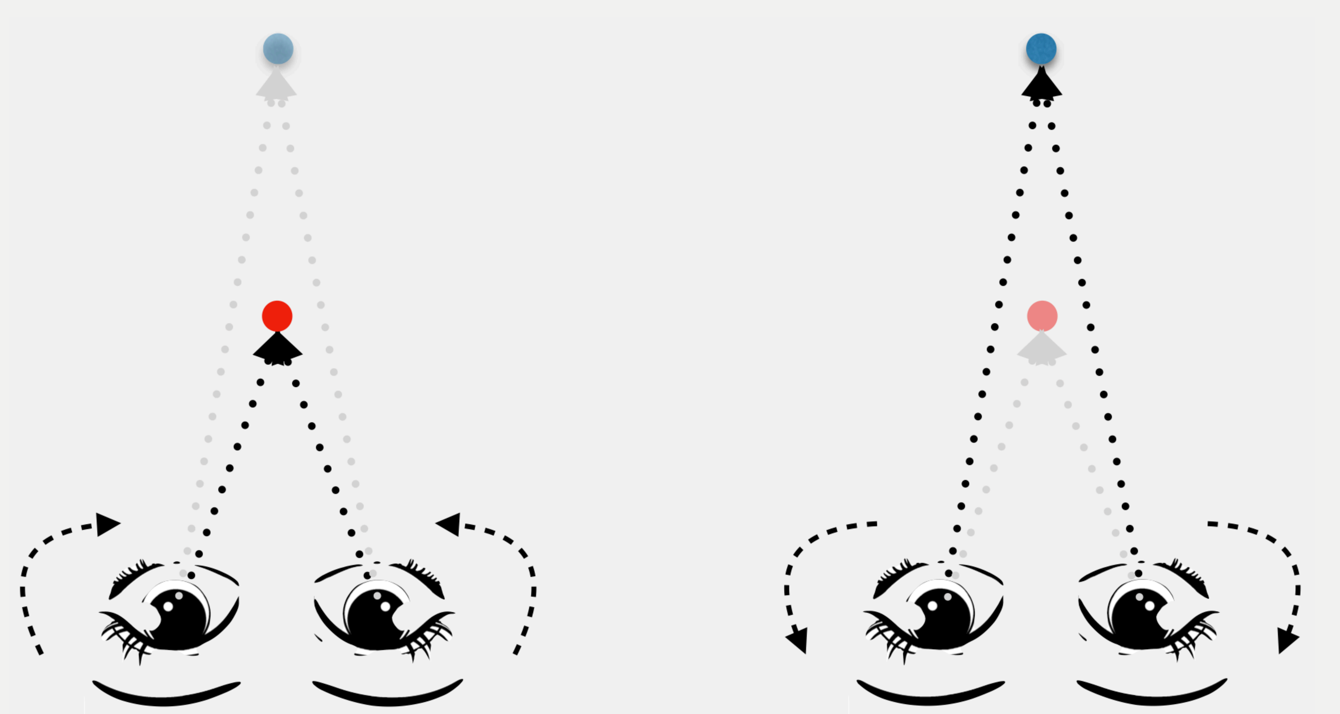


Fig. 2. Human eyeballs rotate inward (outward) when gazing at a near (far) target.

Depth From Gaze

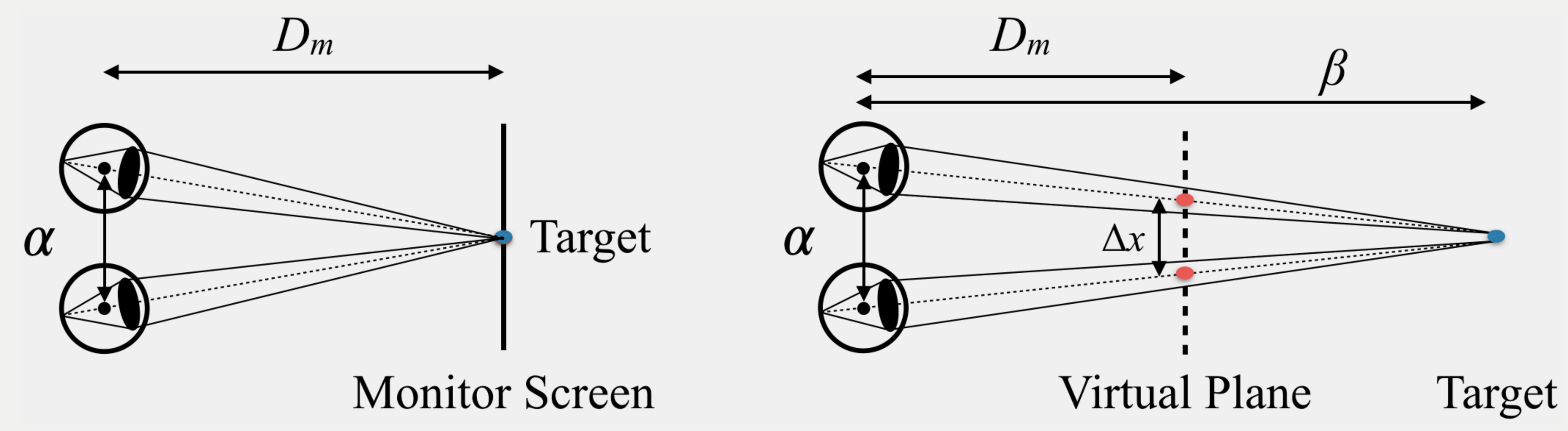


Fig. 3. Top view of 2D and 3D position estimation.

By triangular similarity:

$$\beta = \frac{\alpha D_m}{\alpha - \Delta x} \quad (1)$$

The depth estimate Z for the target is obtained by:

$$Z = \underset{h \in H}{\operatorname{argmax}}\, P(h|D) \quad (2)$$

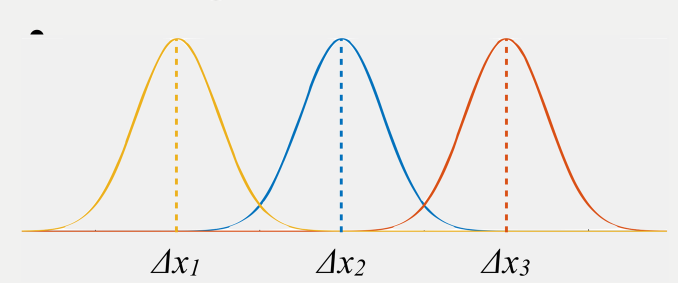
$$Z = \underset{h \in H}{\operatorname{argmin}} \sum_{d \in D} (\Delta x_d - \Delta x_h)^2 \quad (3)$$

D: the set of collected data H: the set of possible depths

Δx_d & Δx_h : the Δx obtained when the target is located at d and h

Minimal Distance In Depth

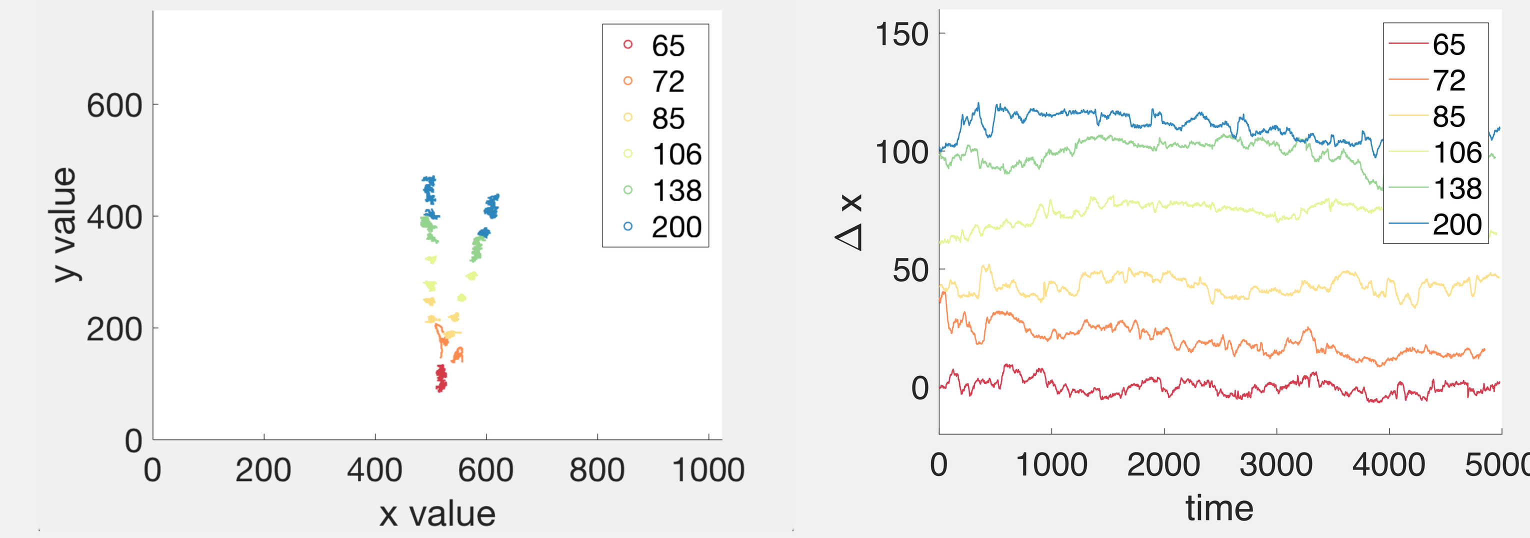
- The convergent sensitivity of human eyes decreases with target distance.
- Consequently, the minimal distance between two distinguishable depths increases with target distance.



$$P = 2\left(1 - F_2\left(\Delta x_2 + {}^{I}/_{2}\right)\right) \quad (4)$$

Example Results

- Fixation eye-movements introduce variation to the visual axes.
- The disparity Δx becomes larger as the target is placed farther.



Evaluation

●: 65 cm; ●: 72 cm; ●: 85 cm; ●: 106 cm; ●: 138 cm; ●: 200 cm.

subject	Proposed Method: best	Proposed Method: 2nd best	Proposed Method: others	Baseline Method: best	Baseline Method: 2nd best	Baseline Method: others
S_1	12	0	0	10	2	0
S_2	13	5	0	3	3	12
S_3	7	11	0	3	5	10
S_4	8	3	1	2	4	6
S_5	11	6	1	3	3	12
S_6	5	7	0	2	2	8
S_7	6	5	1	5	5	2
	61%	36%	3%	27%	24%	49%

- Our method is able to tell the relative depth between target positions.
- The best and the second-best estimates together account for 97% of the test.
- The quality of depth estimate from gaze varies from one subject to another.