UNIVERSITY OF TAMPERE
SIS SCHOOL OF INFORMATION SCIENCES
TAUCHI TAMPERE UNIT FOR COMPUTER-HUMAN INTERACTION

Oleg Špakov
Yulia Gizatdinova

Real-Time Hidden Gaze Point Correction

Motivation

The accuracy of gaze-pointing is limited by a number of technological and biological constrains. Despite that, sometimes the exact and correct point-of-regard can be determined very accurately. We hypothesized that such events could be recorded and used later to reduce the pointing error.

Challenge

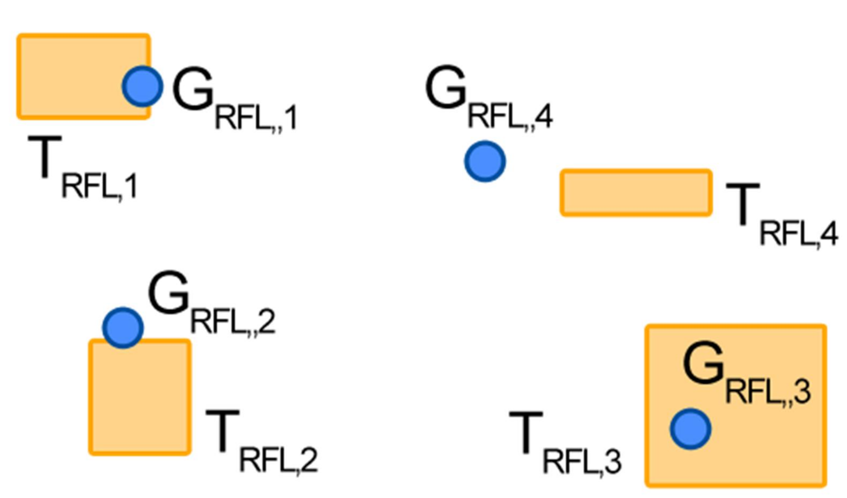
The real-time hidden gaze correction (RTHGC) algorithm extends the use of the implicit required fixation locations (RFL) proposed by Hornoff and Halverson, 2002. It collects the fixations with the highly confident links to objects, and the collected map of disparities between gaze and objects is used further to estimate the probability of each onscreen object being focused given the current gaze point.

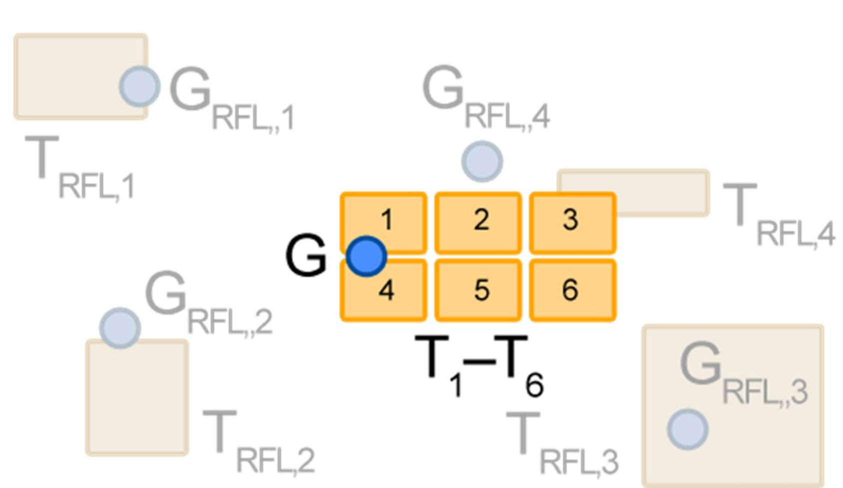
As the objects are not single points but have a certain size, and, in general, all points inside of them have equal probability to be hit by gaze, a sophisticated algorithm was developed to calculate the probability ***P*** for the current gaze-point ***G*** to hit a certain object ***T*** .

Algorithm

Input:

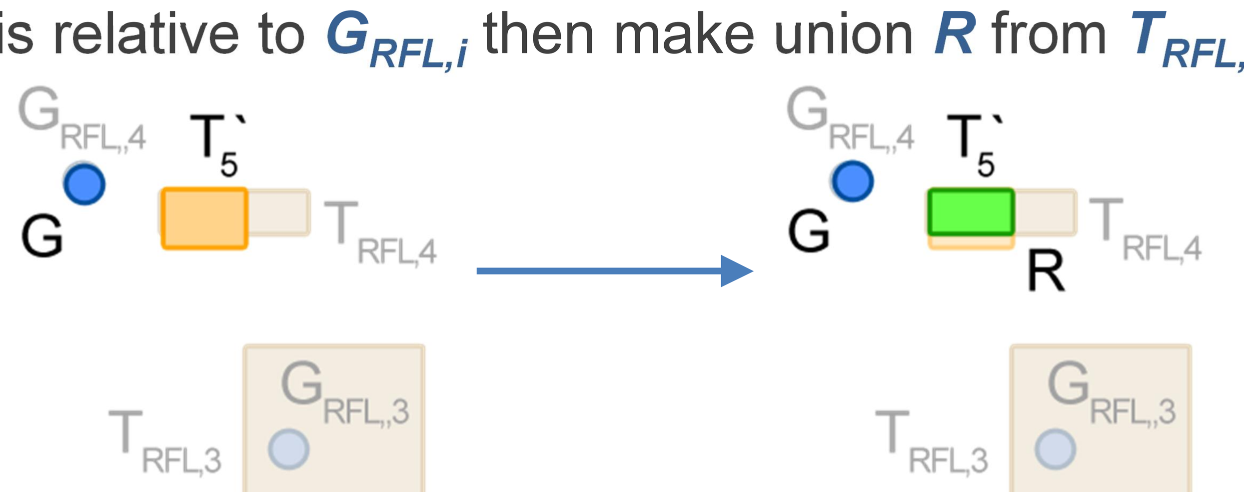
- Database with the confident links between gaze points G_{RFL} and rectangular objects T_{RFL}:
- Current gaze point ***G*** and objects T_1-T_N:





Steps on iteration through every RFL:

- Move each T_i so that it becomes T_i' same way relative to ***G*** as $T_{RFL,i}$ is relative to $G_{RFL,i}$ then make union ***R*** from $T_{RFL,i}$ and T_i' :



- Compute horizontal and vertical components of the probability P_i of the current gaze point ***G*** falling into this intersection :

$$P_{X,i} = \frac{F(R.right) - F(R.left)}{F(T.right) - F(T.left)}, \qquad P_{Y,i} = \frac{F(R.bottom) - F(R.top)}{F(T.bottom) - F(T.top)}$$

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} dx, \qquad \sigma = 50$$

Final *P* applying temporal and spatial weights:

$$P = \frac{\sum_i W_{Xi} \sum_i P_{Xi}}{\sum_i W_{Xi}} \cdot \frac{\sum_i W_{Yi} \sum_i P_{Yi}}{\sum_i W_{Yi}}$$

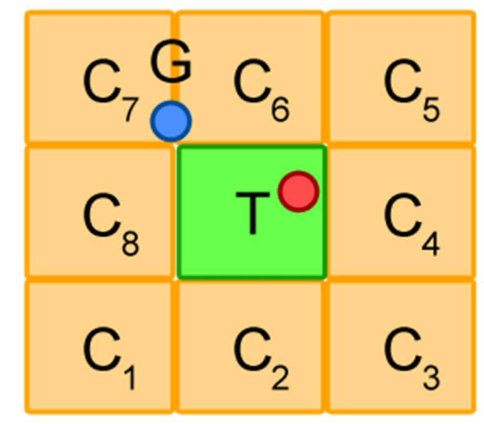
$$W = W_d W_t, \quad d = |G - G_{RFL}|, \quad s = \begin{cases} T_{RFL}.height \\ T_{RFL}.width \end{cases},$$

$$W_x = \begin{cases} e^{-x^2/2\sigma^2}, & x \geq 2\sigma \\ 0 & x < 2\sigma \end{cases}, \quad \sigma_d = 150, \quad \sigma_s = 85$$

Verification

Twelve participants shortly (1.3 seconds) observed small targets (ø = 6 pixels) appearing 10 times randomly in each cell of 4x5 full-screen sized grid. A calibration of Tobii T60 system with 2, 5 or 9 points preceded each of 3 sessions. Each of session consisted of 3 blocks 200 trials each, with some participant movement in between to test the algorithm resistance against the degrading pointing accuracy.

The recorded data was used to simulate observations of 16, 32 …144 pixels targets with 8 surrounding candidates:



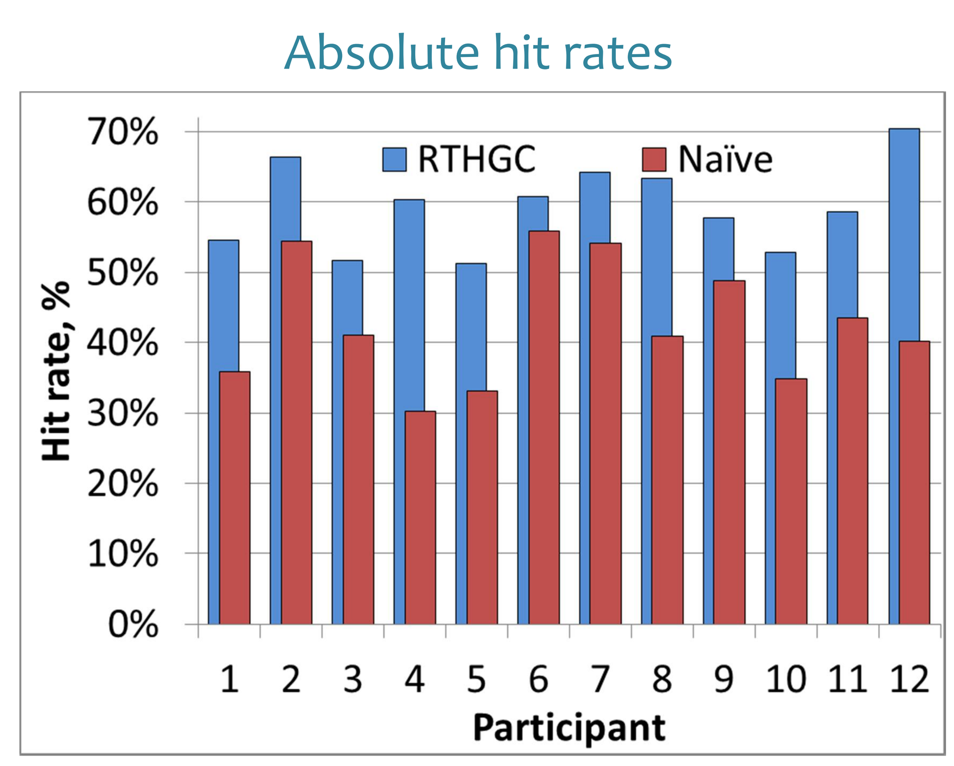
T – simulated target, includes the displayed target (red circle)

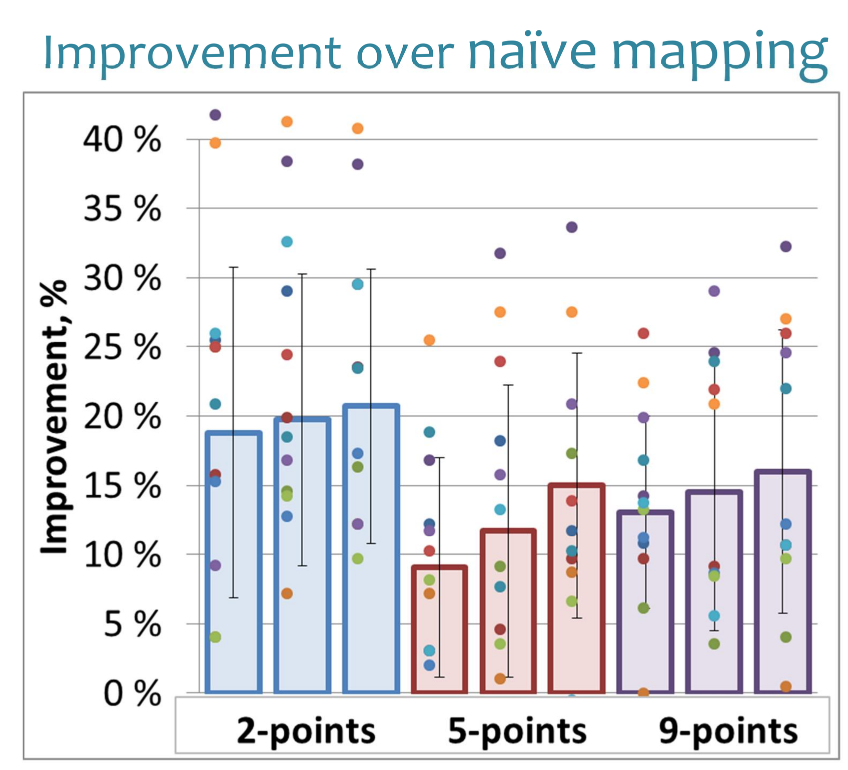
C – simulated candidates

G – the recorded fixation

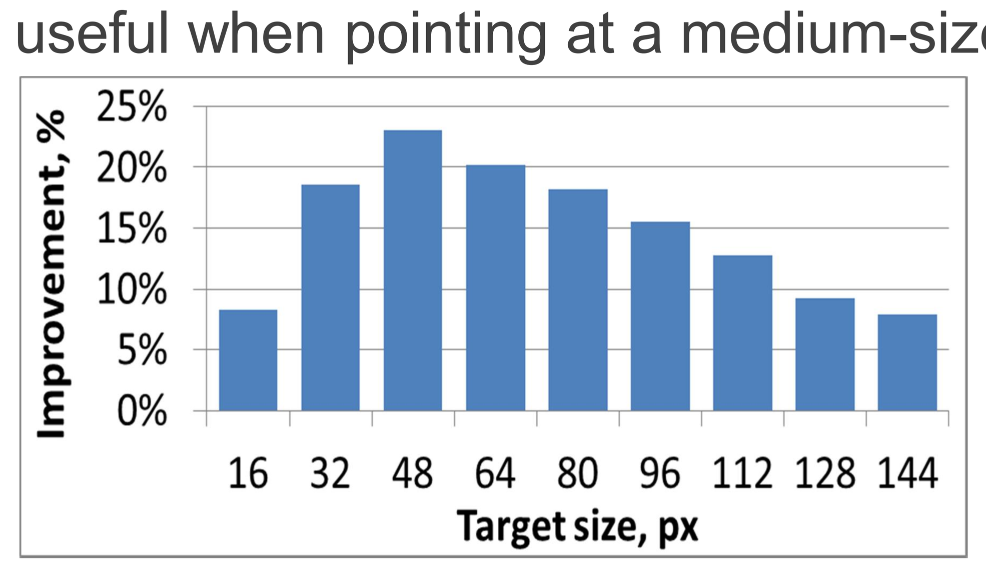
In each simulated trials, the preceded trials were used a list of the confident links between gaze points G_{RFL} and rectangular objects T_{RFL} The verification was completed against the naïve gaze-to-object mapping. Some outlying cases were removed from the analysis (0.25%).

The grand mean rate of hitting a target (HR) using naïve algorithm was μ= = 42% (σ = 5.7%). The proposed algorithm increased the it by 15.7% in average (μ= 57.7%, σ = 5.7%), and by 20% for the last 25 trials. It was beneficial for each participant (at least 4.8%, at most 30%) and in each condition:





The proposed algorithm has appeared indeed resistant against accuracy degradation due to body movements, and it was the most useful when pointing at a medium-sized targets:



Fit:

$$HR_{RTHGC} - HR_N = \frac{0.006x}{1 + (0.01467x)^3}$$

$$R = 0.991$$

$$x = [16 - 144, 16]$$

Conclusions

The algorithm safely improves the gaze-pointing accuracy whenever some fixations can be confidently linked with targets. It is the most efficient for the middle-sized objects: in this study the improvement was by 23.5% on average in comparison to the naïve mapping for targets of the 48 pixels. It works in real-time, is independent of a task and object locations, and does not penalize the performance in term of the increased pointing task. We suppose that this algorithm may work in parallel with other algorithms that improve gaze pointing.