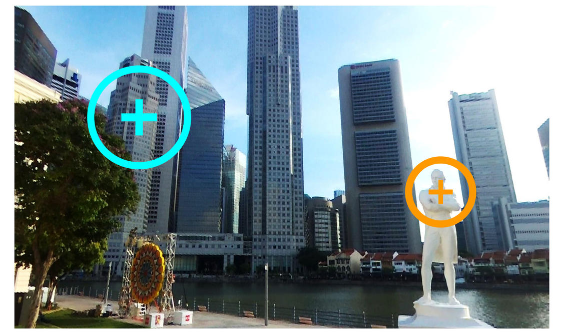
Solving the Perspective-2-Point Problem for Flying-Camera Photo Composition

Ziquan Lan, David Hsu, Gim Hee Lee

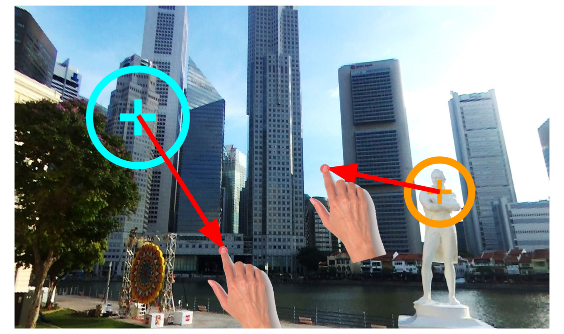
National University of Singapore

INTRODUCTION

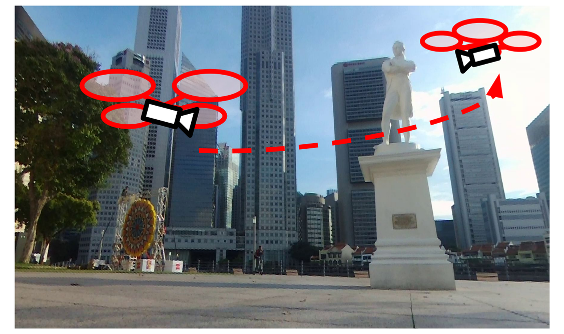
Automatic viewpoint selection for flying-camera photo composition:



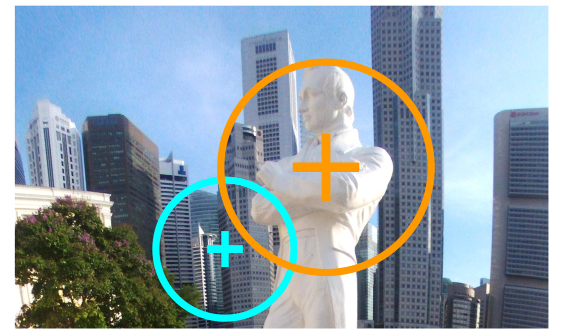
The initial viewpoint with two objects



The user composes a photo with gestures



The camera finds a desired viewpoint



The final photo with desired composition

Proposed solution:

- Formulate the under-determined P2P as a constrained nonlinear optimization
- Solve the constrained nonlinear optimization in *closed form*

CONSTRAINED NONLINEAR OPTIMIZATION

Given:

- $\mathbf{q}_j^W$: estimated object centroid positions in the world frame F_W.
- $\mathbf{p}_j^I$: desired composition positions in the image frame F_I.
- ϵ_j: estimated collision-free distances to the objects.
- $\mathbf{t}_0^W$: the camera's initial position in F_W.
- $\mathbf{K}$: the camera's intrinsic matrix.

Find: $j = 1,2$

The *nearest* camera pose $(\mathbf{R}_C^W, \mathbf{t}_C^W)$ from $\mathbf{t}_0^W$

$$\underset{\mathbf{R}_C^W,\, \mathbf{t}_C^W}{\mathrm{argmin}} \left\| \mathbf{t}_C^W - \mathbf{t}_0^W \right\|^2, \tag{1}$$

subject to

Two 2D-3D point correspondences

$$\lambda_j \mathbf{p}_j^I = \mathbf{K}\left(\mathbf{R}_W^C\, \mathbf{q}_j^W + \mathbf{t}_W^C\right), \tag{2}$$

Avoid collision with the two objects

$$\left\| \mathbf{t}_C^W - \mathbf{q}_j^W \right\| \geq \epsilon_j. \tag{3}$$

P2P SOLUTION IN CLOSED FORM

- **Construct an auxiliary frame F_A:** one axis passes through $\mathbf{q}_1^W$ and $\mathbf{q}_2^W$.

$$\mathbf{q}_1^A = [0 \quad 0 \quad 0]^T, \mathbf{q}_2^A = \left[\left\|\mathbf{q}_1^W - \mathbf{q}_2^W\right\| \quad 0 \quad 0\right]^T.$$

- **Reformulate the equality constraints Eq.** (2) **using F_A:**

$$\lambda_j \mathbf{p}_j^I = \mathbf{K}\left(\mathbf{R}_A^C\, \mathbf{q}_j^A + \mathbf{t}_A^C\right). \tag{4}$$

  - Rewrite Eq. (4) into

$$\mathbf{A}\,\mathbf{w} = \mathbf{0}_{4\times 1},$$

  - $\mathbf{A}$ is a 4x6 matrix from $\mathbf{K}^{-1}$, $\mathbf{p}_j^I$ and $\left\|\mathbf{q}_1^W - \mathbf{q}_2^W\right\|$,
  - $\mathbf{w} = \begin{bmatrix} \mathbf{c_1}_A^C \\ \mathbf{t}_A^C \end{bmatrix}$, where $\mathbf{c_1}_A^C$ is the 1st column of $\mathbf{R}_A^C = \left[\mathbf{c_1}_A^C \quad \mathbf{c_2}_A^C \quad \mathbf{c_3}_A^C\right]$,
  - Null space of $\mathbf{A}$ gives $\mathbf{w} = \alpha_1 \boldsymbol{e}_1 + \alpha_2 \boldsymbol{e}_2$, (2 parameters: α_1 and α_2),
  - Since $\left\|\mathbf{c_1}_A^C\right\| = 1$, both $\mathbf{c_1}_A^C$ and $\mathbf{t}_A^C$ are parameterized with 1 parameter θ.

- **Reformulate the inequality constraints Eq.** (3) **using F_A:**

$$\left\|\mathbf{t}_C^A - \mathbf{q}_j^A\right\| \geq \epsilon_j, \tag{5}$$

  - $\mathbf{t}_C^A = -\left[\mathbf{c_1}_A^C \quad \mathbf{c_2}_A^C \quad \mathbf{c_3}_A^C\right]^T \mathbf{t}_A^C = [x \quad y \quad z]^T$,
  - Since $y^2 + z^2 = \left\|\mathbf{t}_A^C\right\|^2 - x^2$, we introduce 1 more parameter ϕ so that y and z are parameterized with θ and ϕ.
  - Plugging $\mathbf{t}_C^A$ in Eq. (5), we can solve for eight boundary solutions. ●

- **Reformulate the objective function Eq.** (1) **using F_A:**

$$\underset{\mathbf{R}_C^W,\, \mathbf{t}_C^W}{\mathrm{argmin}} \left\|\mathbf{t}_C^A - \mathbf{t}_0^A\right\|^2, \tag{6}$$

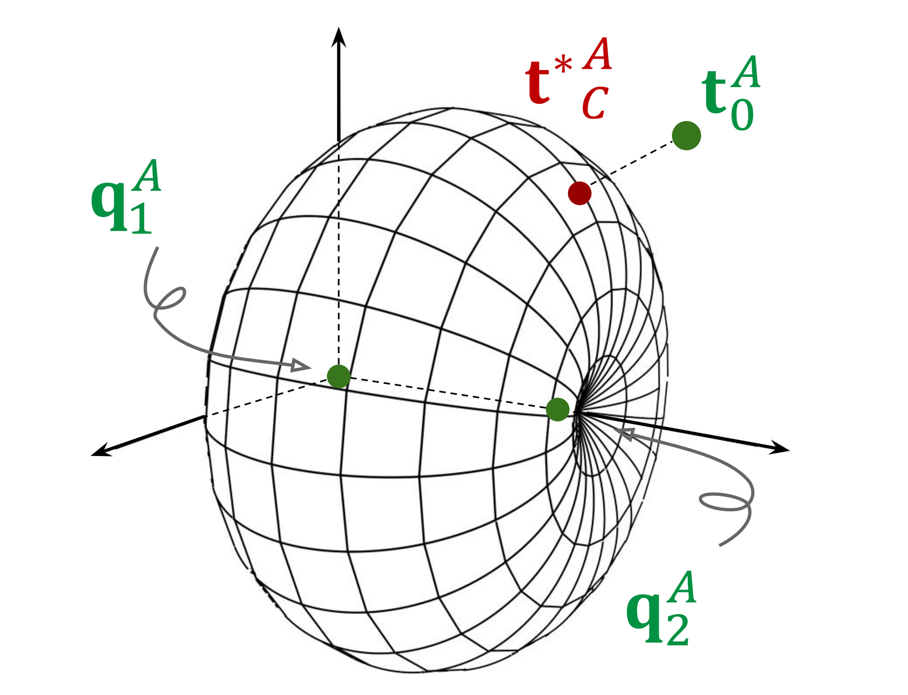
  - Note the objective function $obj = \left\|\mathbf{t}_C^A - \mathbf{t}_0^A\right\|^2$ is parameterized using θ and ϕ.

- **Minimize obj with respect to ϕ and θ:**
  - $\frac{\partial obj}{\partial \phi} = \frac{\partial obj}{\partial \theta} = 0$ yields eight general solutions. ●●

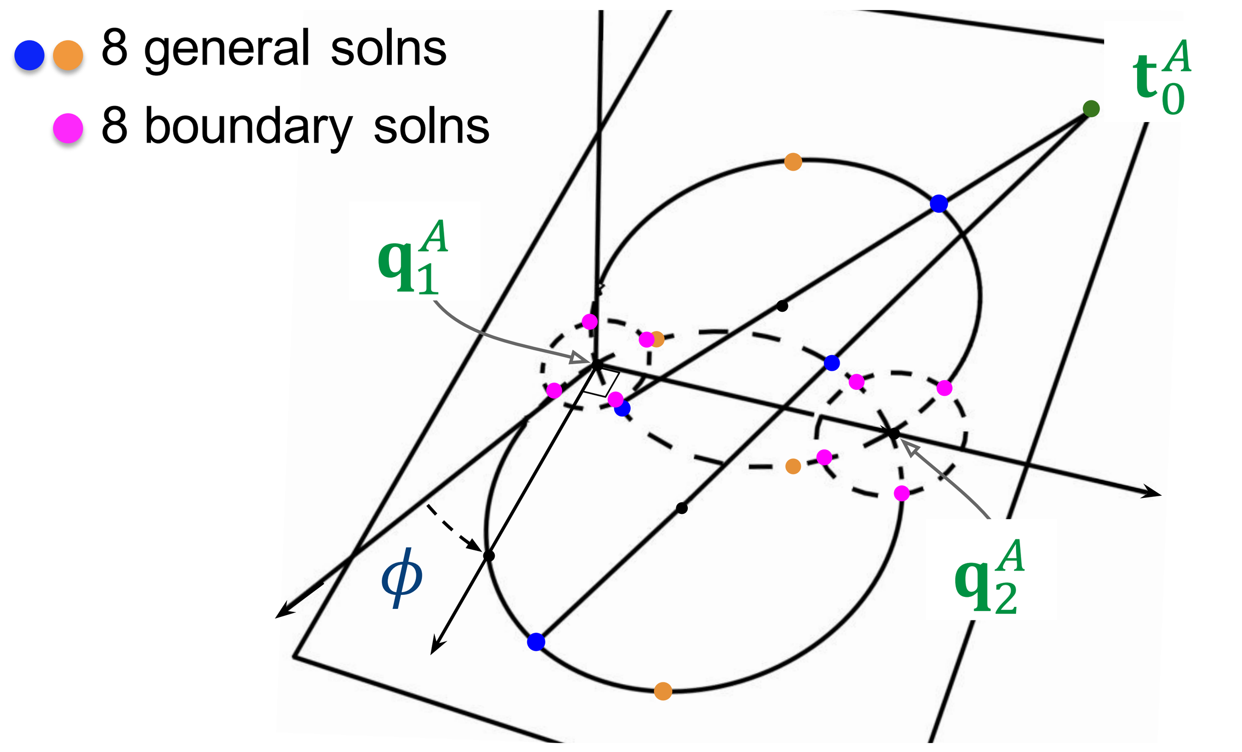
- **Obtain the global optimum** $\mathbf{t}^{*A}_C$, hence $\mathbf{t}^{*W}_C$, by enumeration, then the unique $\mathbf{R}^{*W}_C$.

SOLUTION SPACE

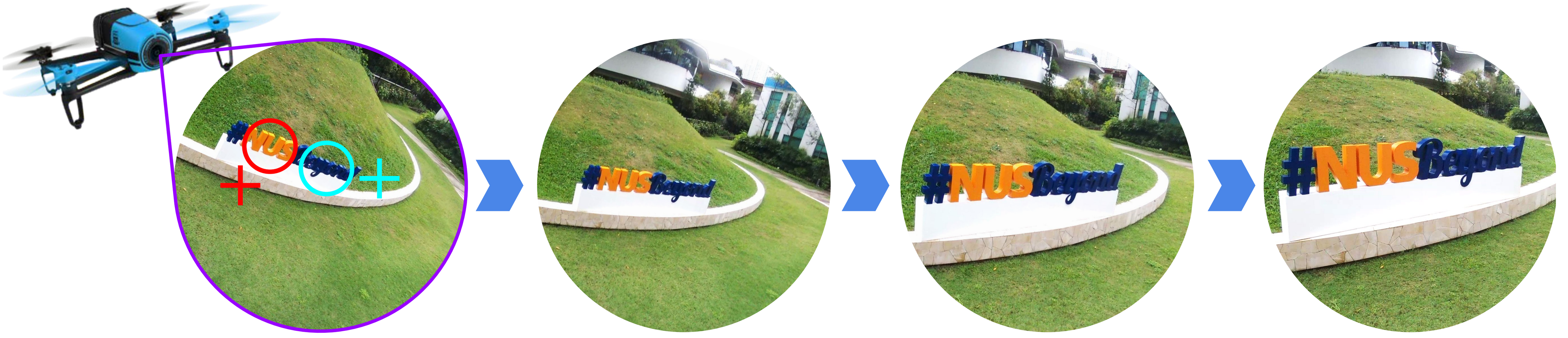
Auxiliary frame F_A



Optimum candidates in F_A

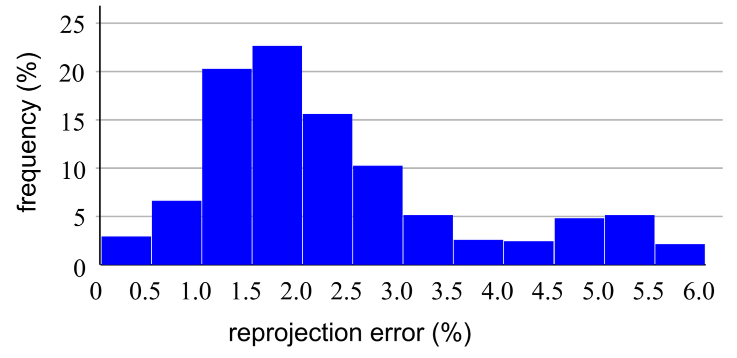


RESULTS



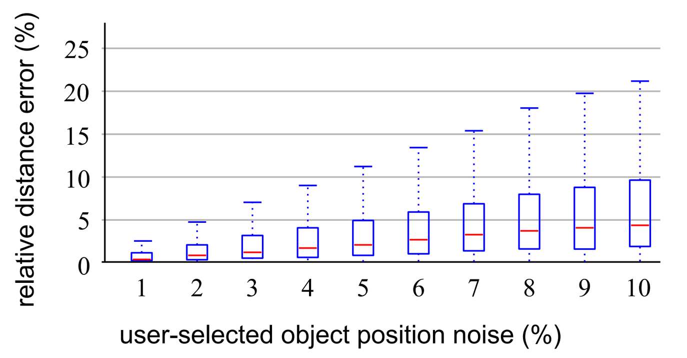
Parrot Bebop drone:

real photo-taking scenario



Synthetic data:

noisy object projections $\mathbf{q}_j^W$



noisy object positions $\mathbf{p}_j^I$