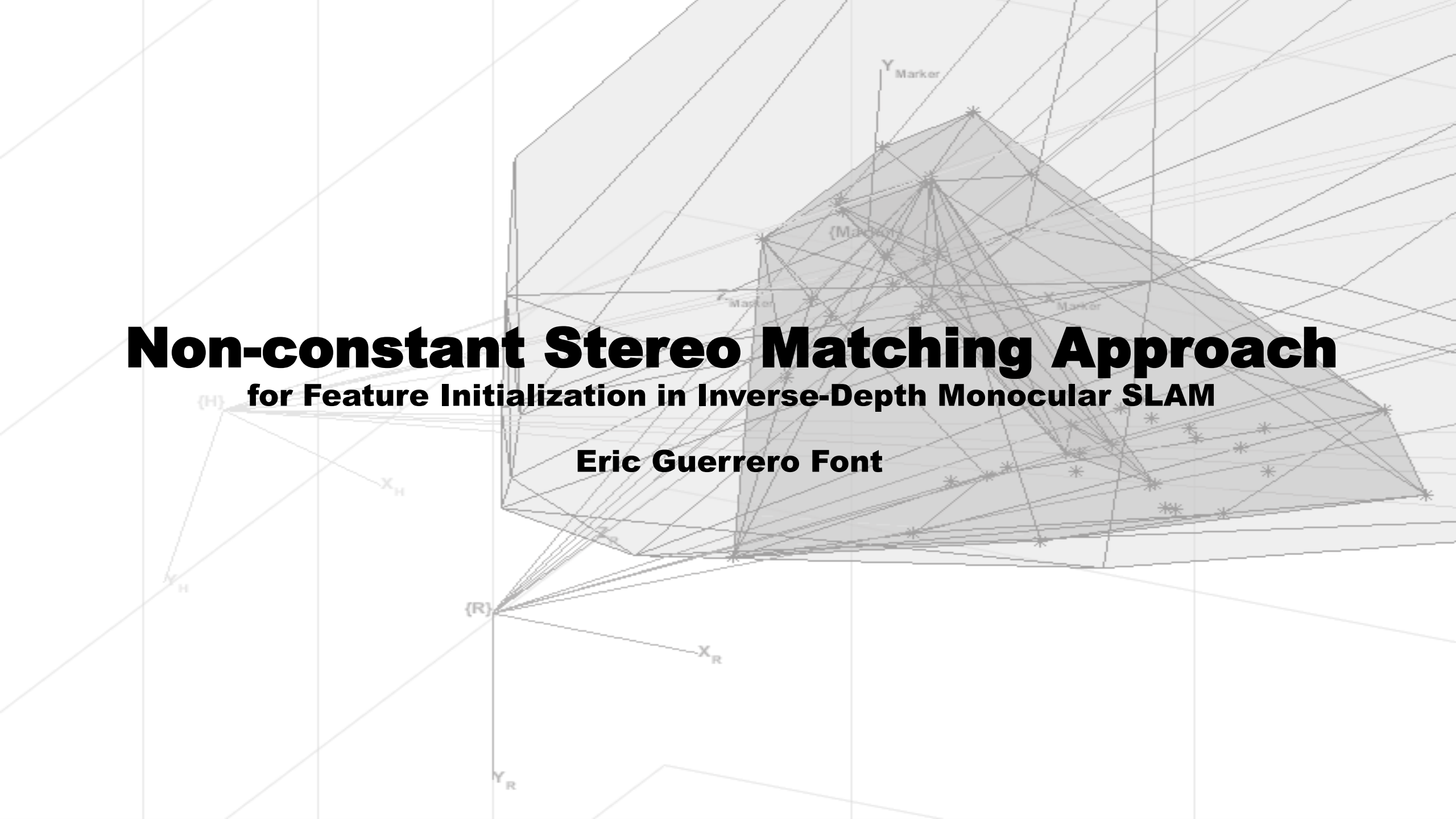


Non-constant Stereo Matching Approach

for Feature Initialization in Inverse-Depth Monocular SLAM

Eric Guerrero Font

Non-constant Stereo Matching Approach

for Feature Initialization in Inverse-Depth Monocular SLAM

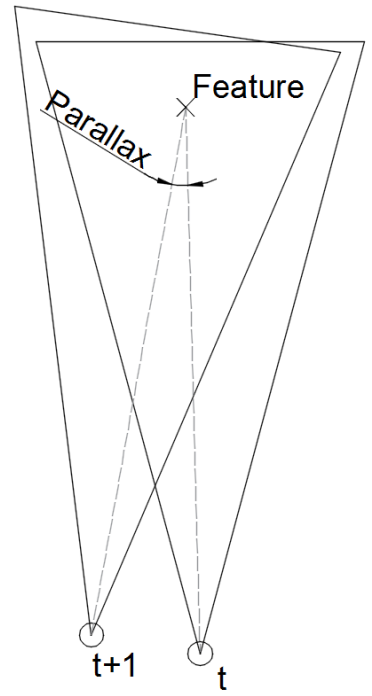
- *Not required to be done for each new frame*

- *Refers to the problem of assigning pixels of two different images that belong to the same part of a 3D scene*

- *Refers to when the map features are going to be use to localize the robot.*
- *More parallax, less uncertainty.*

- *Deal with the EKF assumption of the Gaussian PDF*

- *One robot with one camera.*
- *Estimation of the robot pose and map features.*

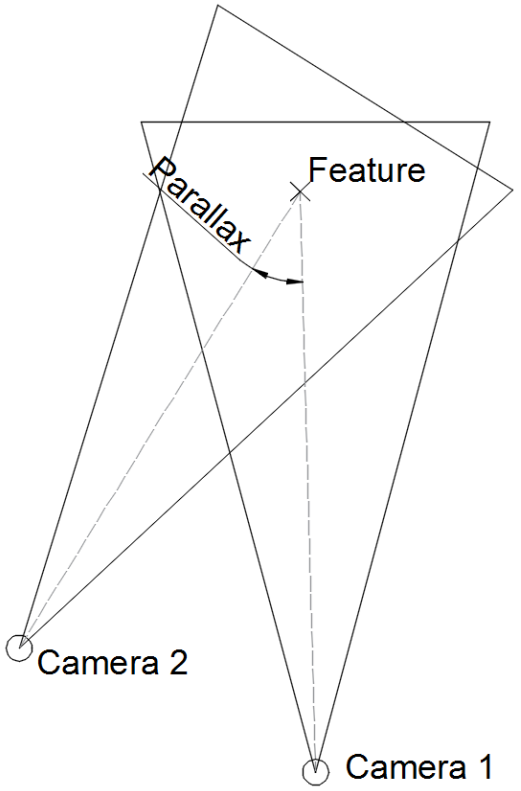


References

E. Guerra, R. Munguia and A.Grau

- *Provide a second camera worn by a human, with known pose.*

- *More parallax, less uncertainty.*
- *Non-constant initialization. When the cameras have a common match, the feature position is estimated.*



Possible Uses

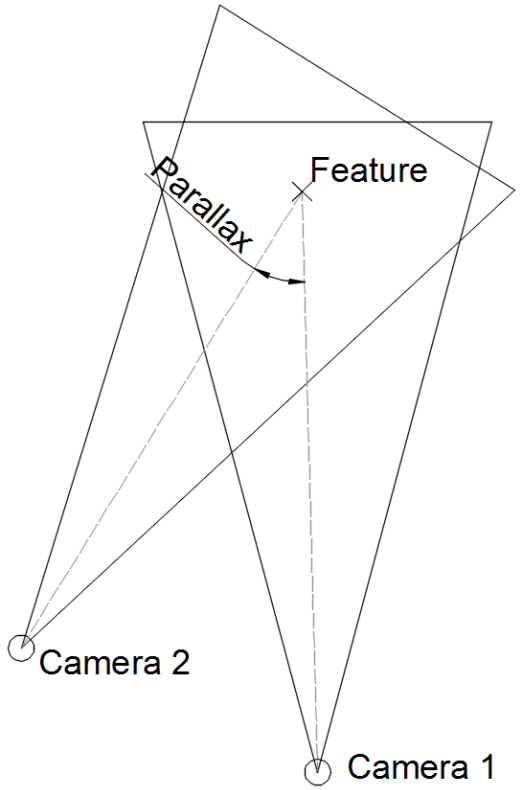
- *Human Robot Interaction.*
- *Cooperative Robots.*

Objectives

Two cameras framework for feature initialization

1- *Provide a tool to know when the matching between images has to be computed. (Avoid outliers and unnecessary consumption of computational resources.)*

2- *Reduce the search space where the match could be.*



Methodology

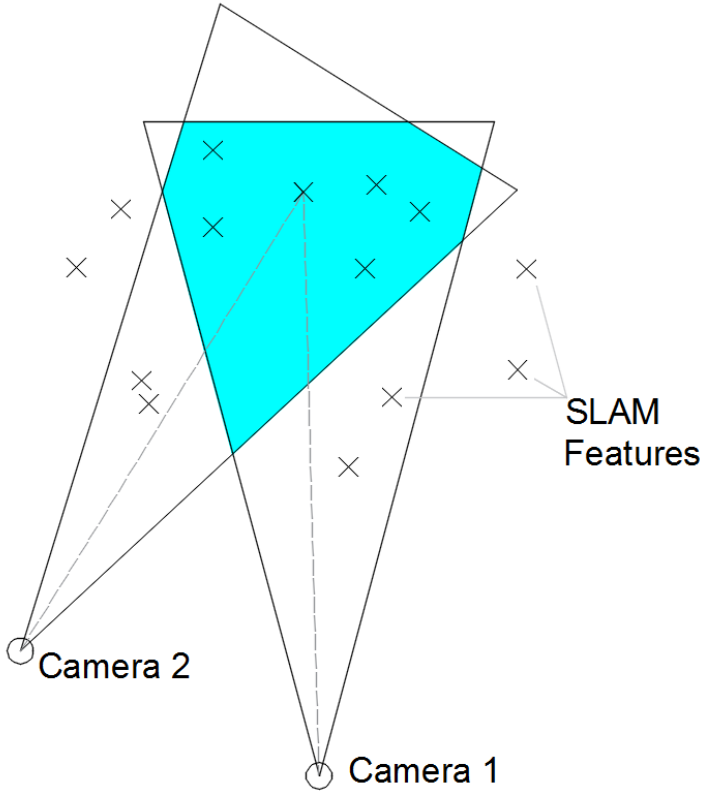
The Common Region

Key facts

- *Common features are inside a common region.*
- *The SLAM algorithm provides actual estimation feature position and covariance.*
- *The pose of the second camera is known. Global reference frame in first camera.*

Goal

- *Reduce the search space for matching.*
- *Get new pose estimation for listed features.*



Methodology

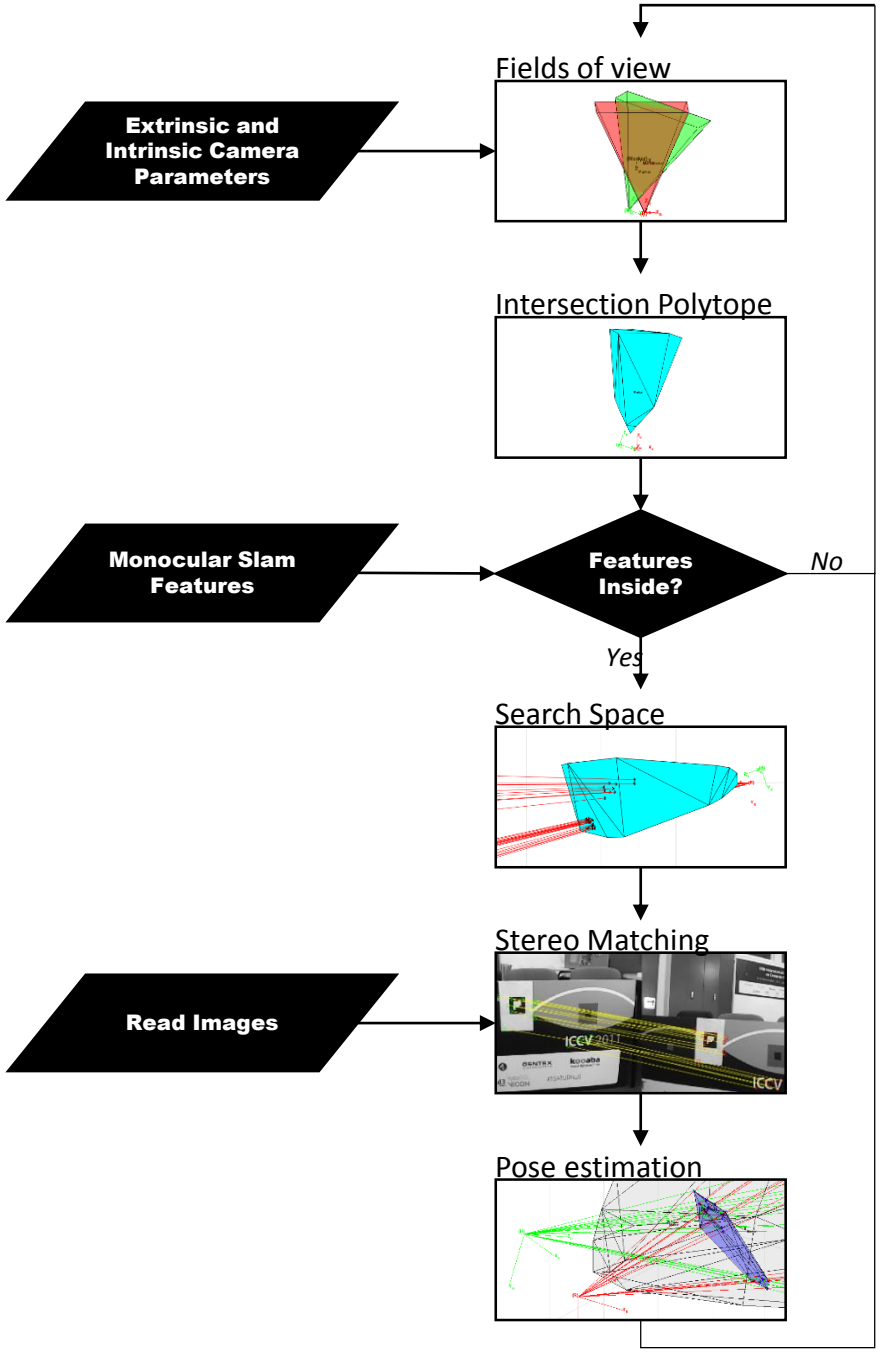
The Common Region

Key facts

- *Common features are inside a common region.*
- *The SLAM algorithm provides actual estimation feature position and covariance.*
- *The pose of the second camera is known. Global reference frame in first camera.*

Goal

- *Reduce the search space for matching.*
- *Get new pose estimation for listed features.*



Methodology

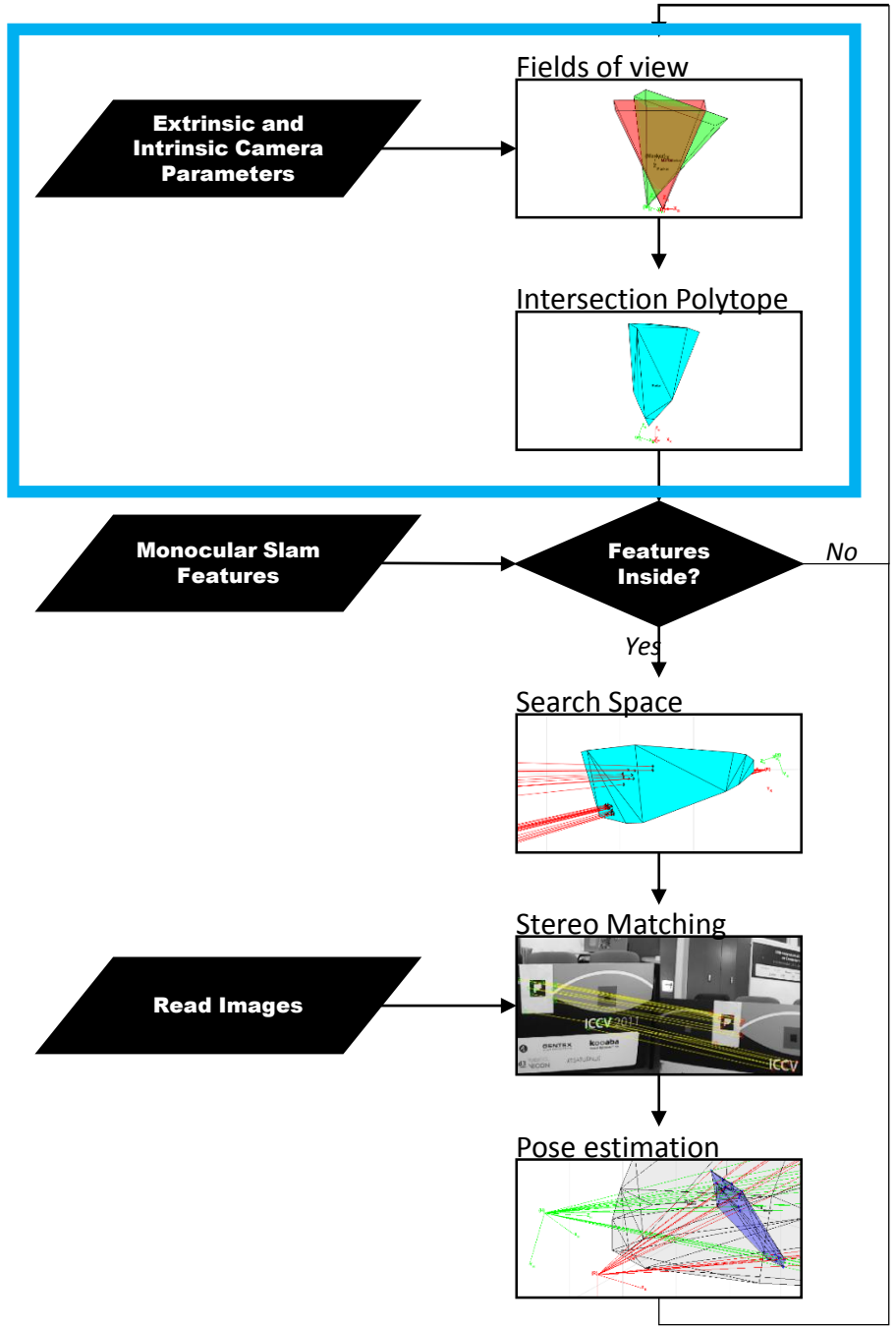
The Common Region

Key facts

- *Common features are inside a common region.*
- *The SLAM algorithm provides actual estimation feature position and covariance.*
- *The pose of the second camera is known. Global reference frame in first camera.*

Goal

- *Reduce the search space for matching.*
- *Get new pose estimation for listed features.*



Fields of view

As pyramids

Compute vertices

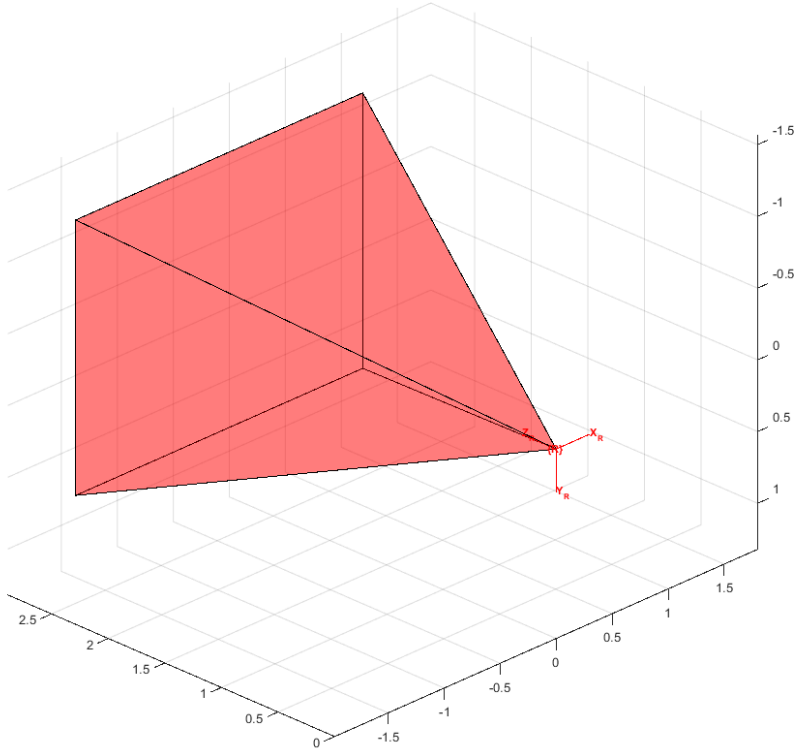
- *Depth given by required minimum parallax.*

$$d_{max} = \frac{\|p_{c1} - p_{c2}\|_2}{2\ tg(\frac{\varphi_{min}}{2})}$$

- *Angular from intrinsic parameters.*

$$[X\ \ Y\ \ Z\ \ 1]^T = P^{-1}\ [u\ \ v\ \ d_{max}\ \ ]^T$$

- *Pose from sensors.*



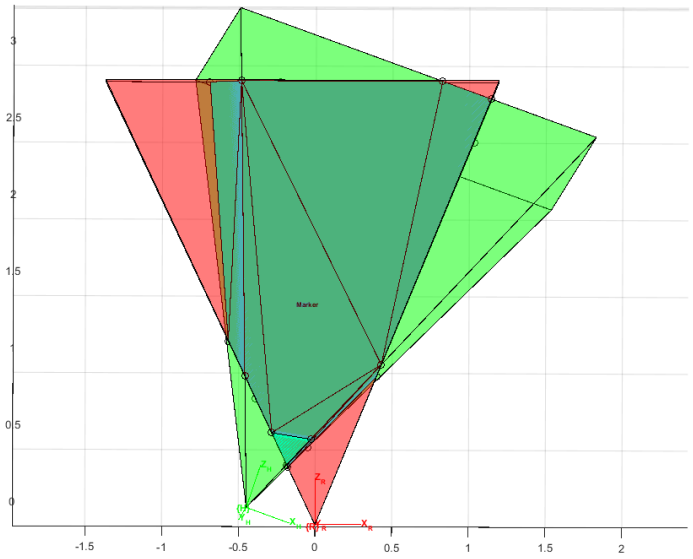
3D field of view

Intersection Polytope

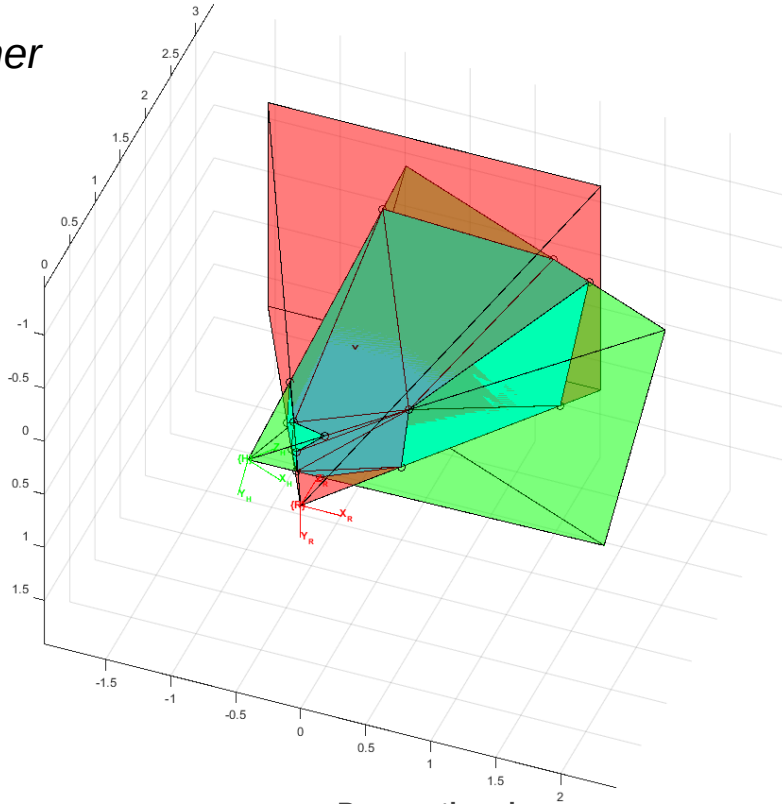
The Common Region

Vertices of the Polytope

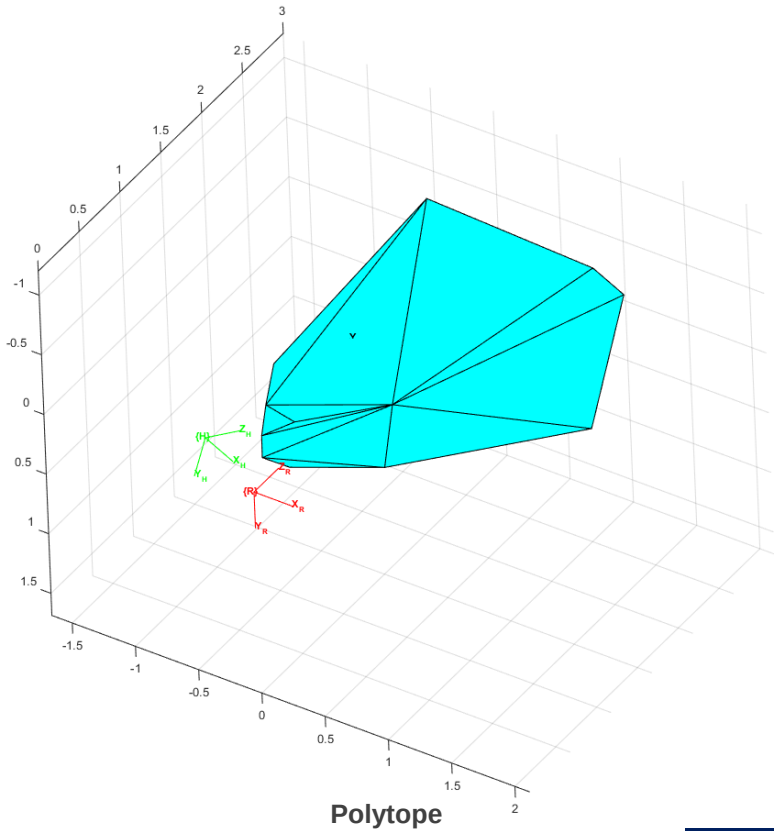
- *Intersection Segments/Planes*
- *Pyramid vertices inside each other*



Top view



Perspective view



Polytope

Methodology

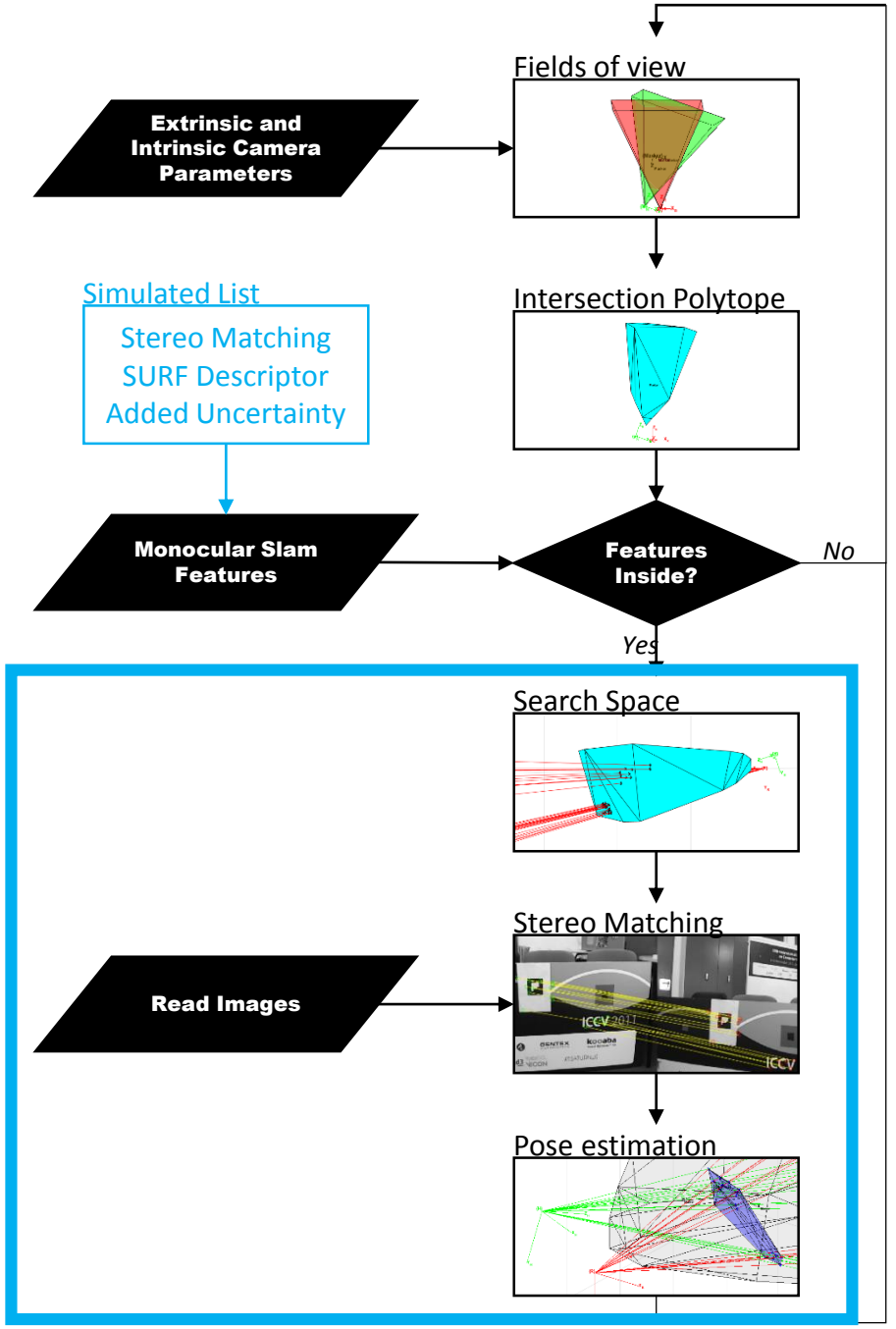
The Common Region

Key facts

- *Common features are inside a common region.*
- *The SLAM algorithm provides actual estimation feature position and covariance.*
- *The pose of the second camera is known. Global reference frame in first camera.*

Goal

- *Reduce the search space for matching.*
- *Get new pose estimation for listed features.*

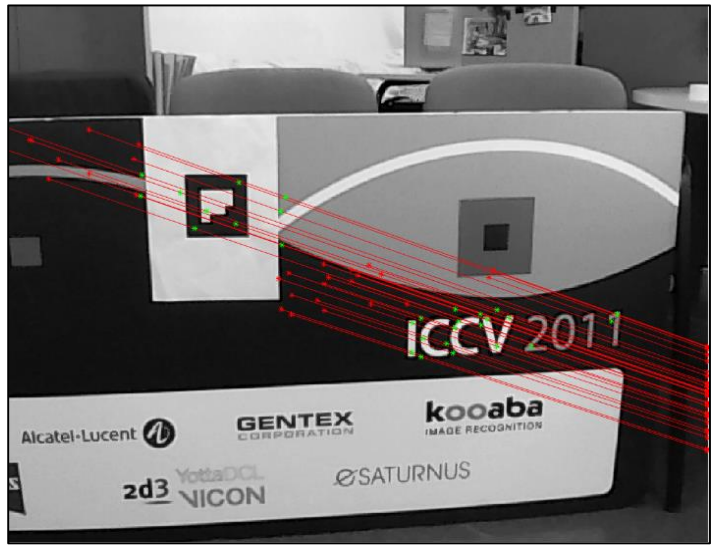


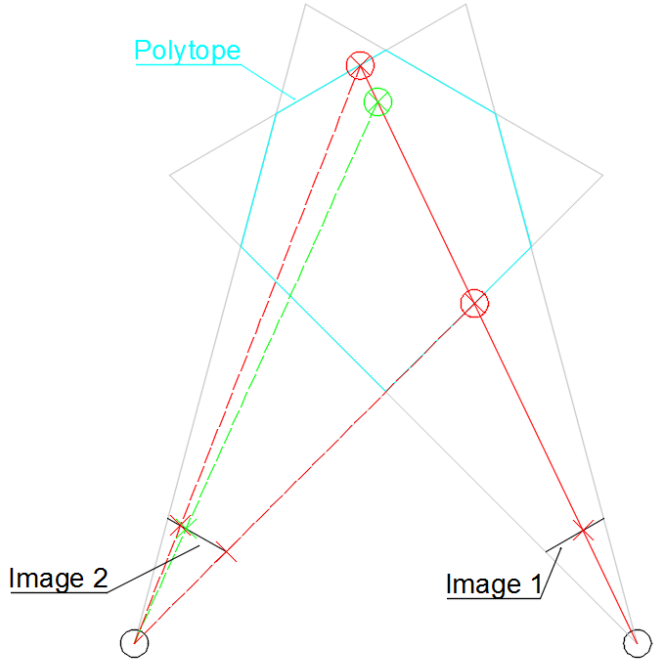
Stereo Matching

Search Space

Polytope Approach

- *Assumption: Uncertainty only in depth.*
- *Project ray from camera 1 to actual feature.*
- *Get intersections with polytope.*
- *Project intersection points to image 2.*
- *The line that joins the pair is the search space.*



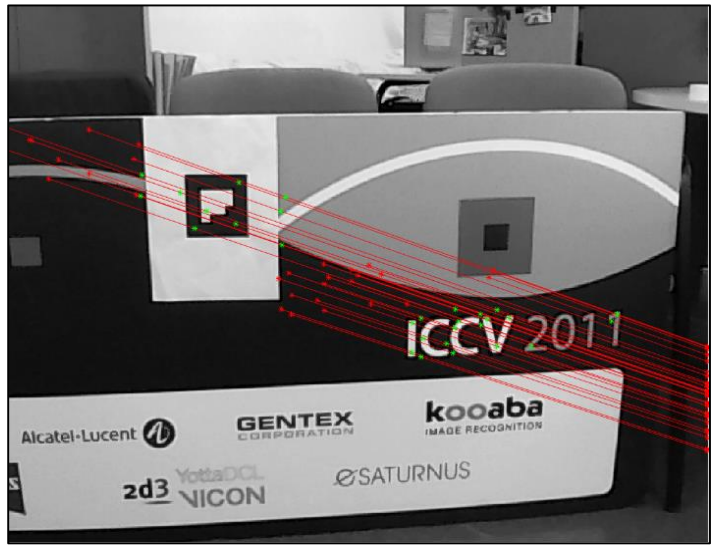


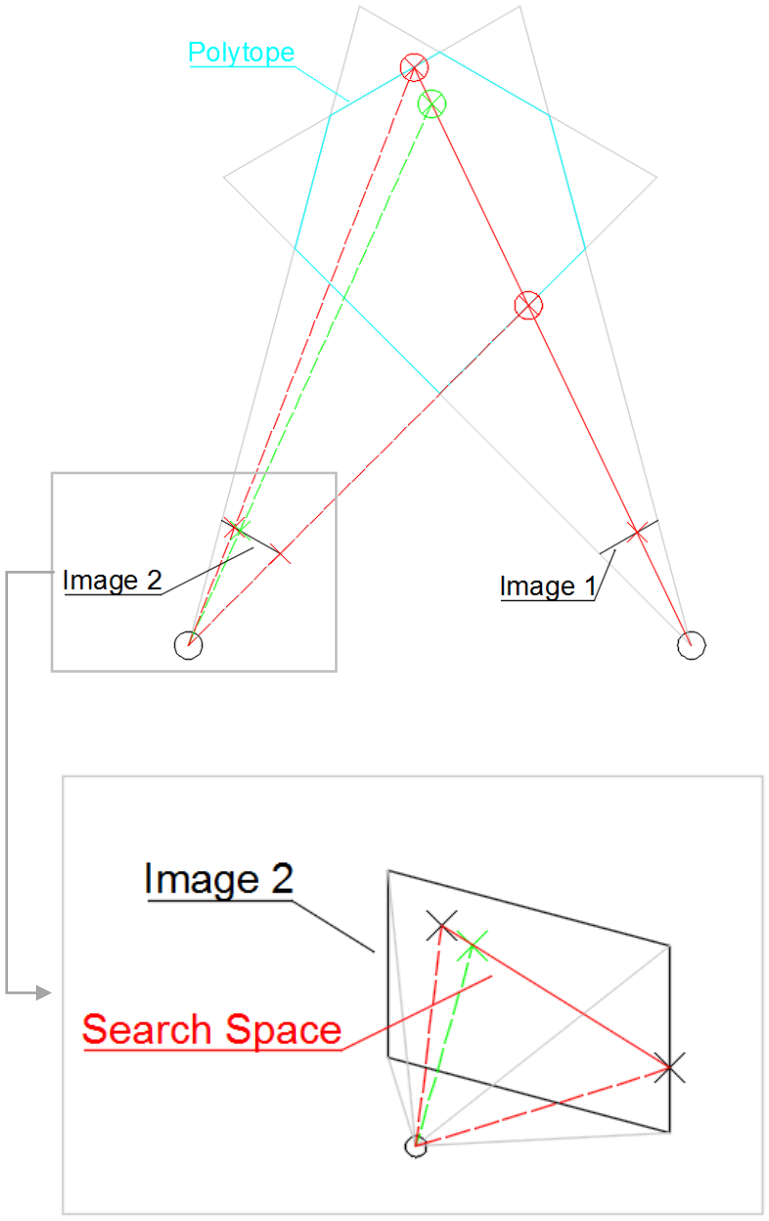
Stereo Matching

Search Space

Polytope Approach

- *Assumption: Uncertainty only in depth.*
- *Project ray from camera 1 to actual feature.*
- *Get intersections with polytope.*
- *Project intersection points to image 2.*
- *The line that joins the pair is the search space.*



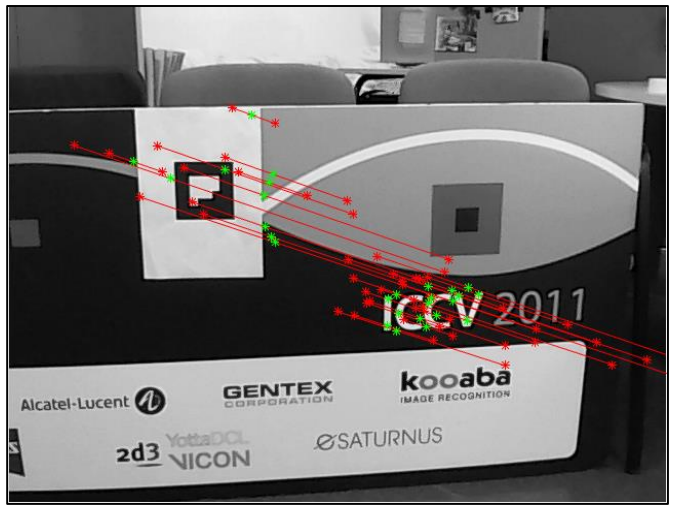


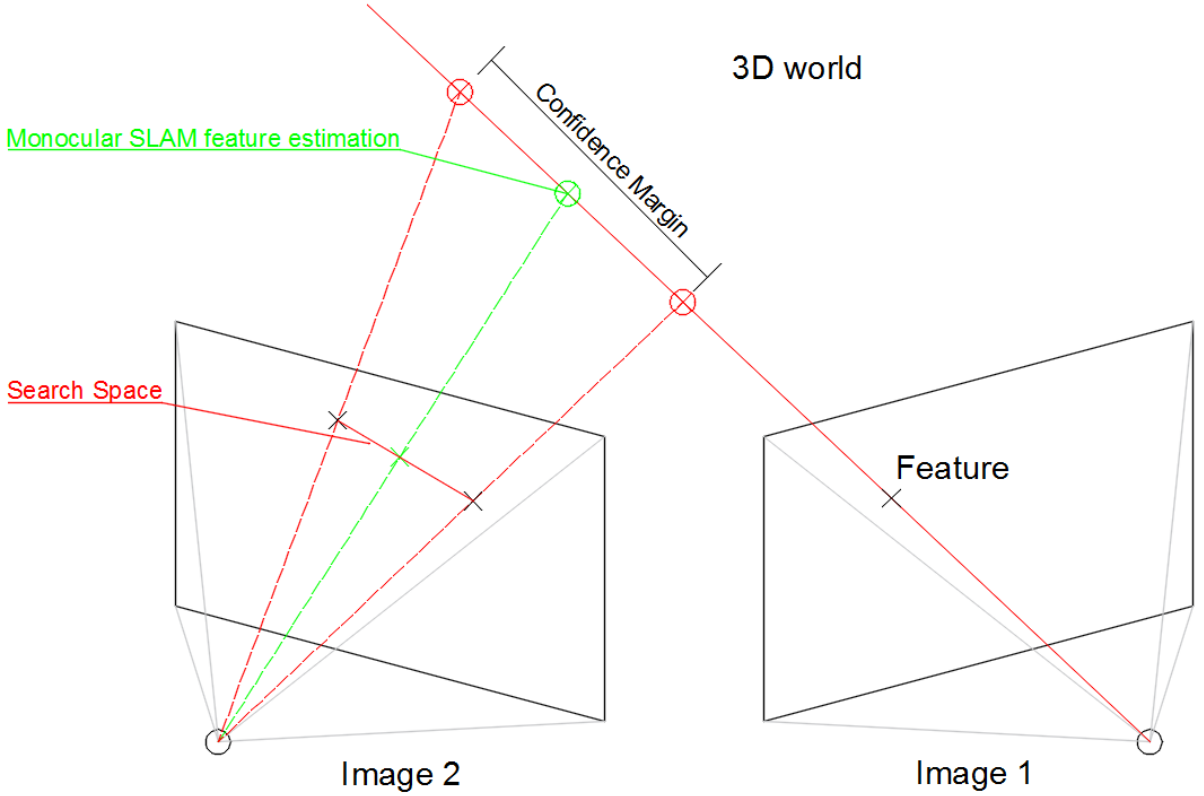
Stereo Matching

Search Space

Covariance Approach

- *Assumption: Uncertainty only in depth.*
- *Project ray from camera 1 to actual feature.*
- *Compute confidence margin with the known feature covariance.*
- *Project margin extremes to image 2.*
- *The line that joins the pair is the search space.*



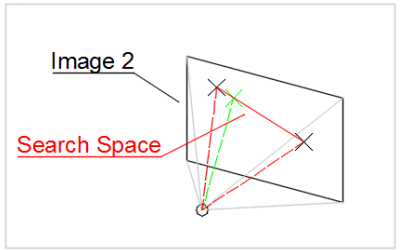


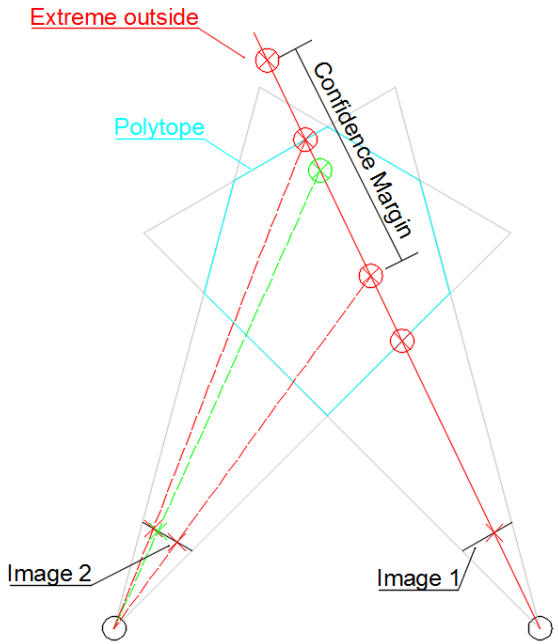
Stereo Matching

Search Space

Final Approach

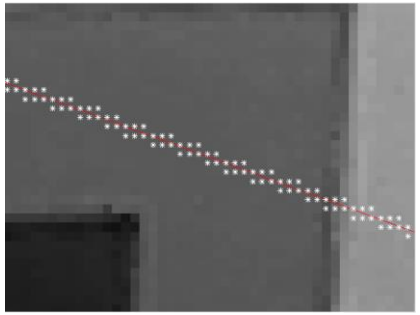
- *Covariance extremes if located inside polytope, if not, intersection ray-polytope.*





Discretization

$$|Mx + N - y| < \varepsilon \quad where \begin{cases} x, y \\ M, N \\ \varepsilon \end{cases} \begin{matrix} Pixel\ Index \\ Line\ Parameters \\ Threshold\ (1\ pixel) \end{matrix}$$



Discretized line.

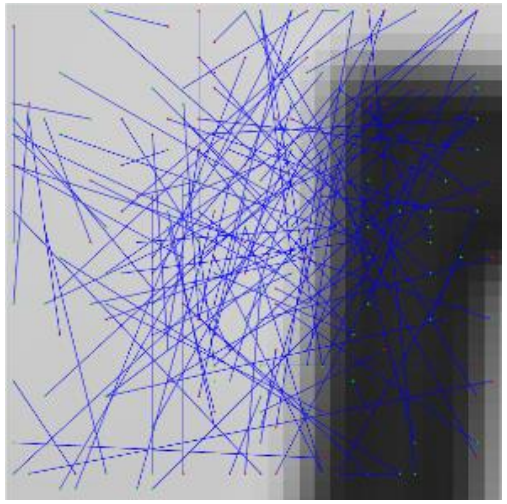
Stereo Matching

BRIEF descriptor

Implementation

- *Smooth the image with a Gaussian filter.*
- *Predefine a random list pairs of points.*
- *The pairs follow a Gaussian distribution* $(0, \frac{1}{10}S^2)$.
- *Descriptor for each patch* ***p*** *as:*

$$\tau(\mathrm{p};\mathrm{x},\mathrm{y}) = \begin{cases} 1 & if \;\; p(x) < p(y) \\ 0 & otherwise \end{cases}$$



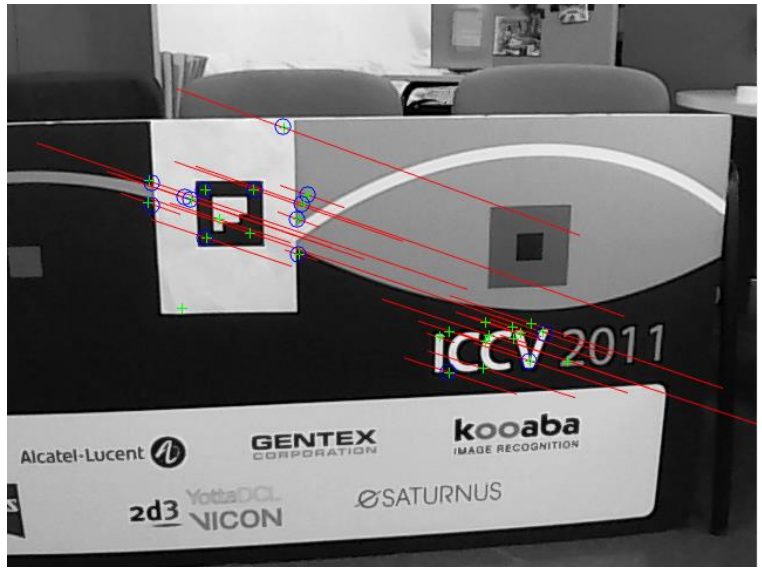
BRIEF descriptor Patch.

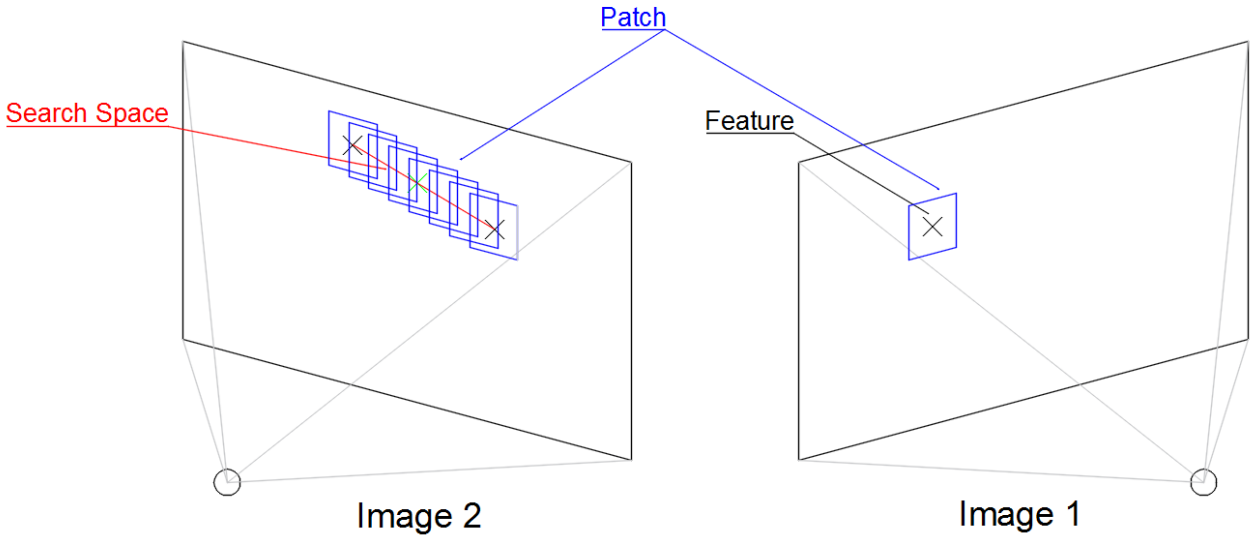
Stereo Matching

BRIEF descriptor

Match Feature/Search Space

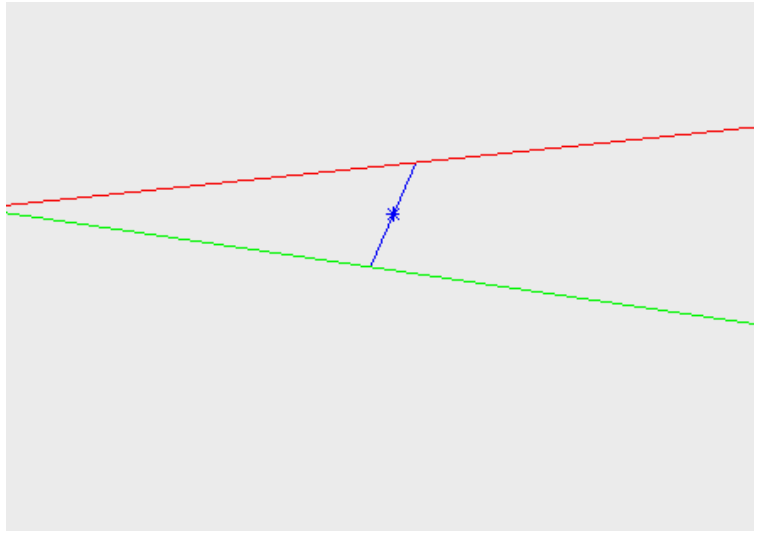
- *Define the BRIEF descriptor in the given position of image 1.*
- *Rotate descriptor and match in the search space of image 2.*
- *Norm-1 to compute the cost.*
- *Threshold on the cost.*



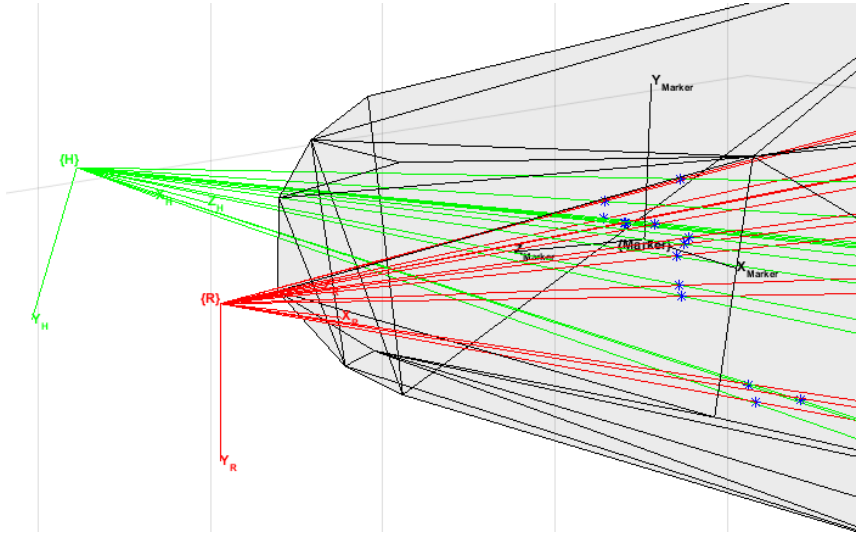


Pose Estimation

- *Project matched pairs.*
- *Find feature position as the pair of rays intersection.*



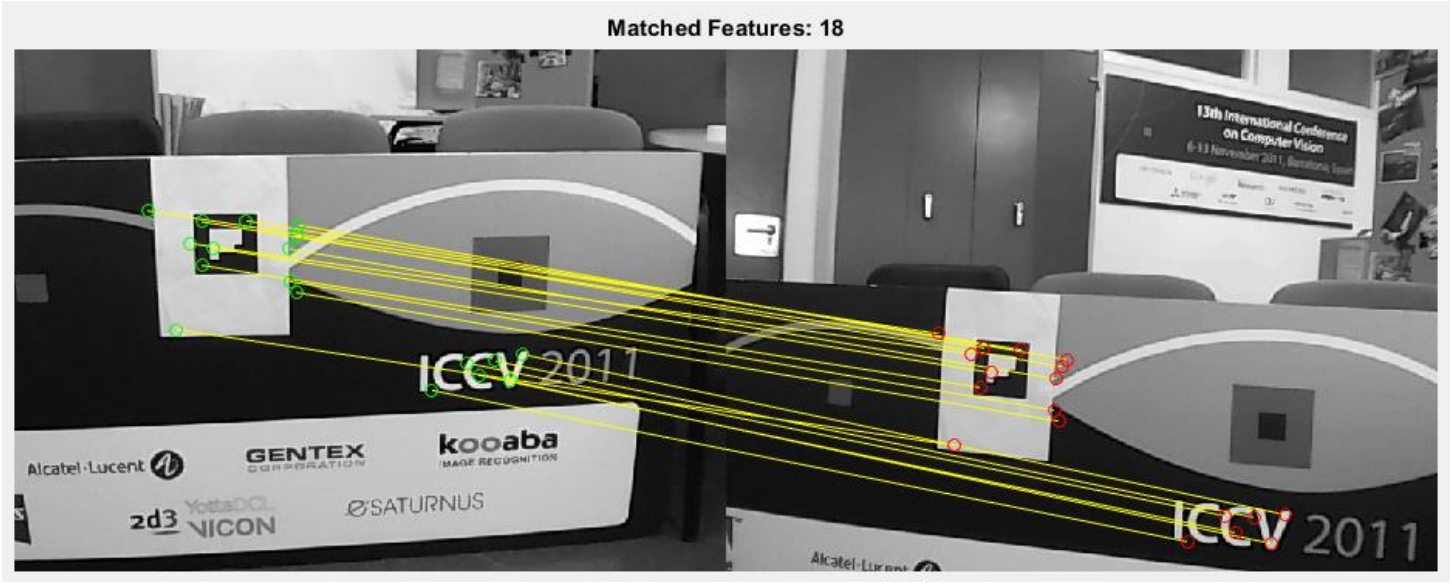
Rays and feature position.



Feature position

Results(I)

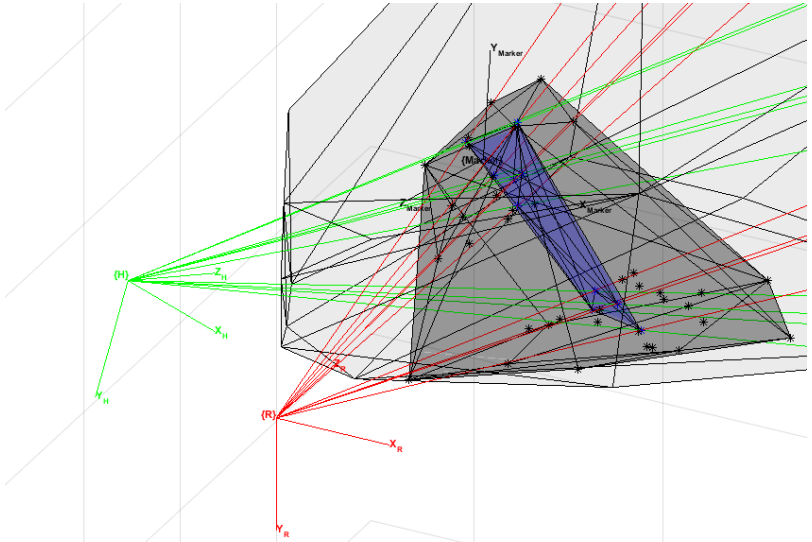
- *Original list of 30 features.*
- *12 features discarded for being outside of the intersection or by the cost threshold in the matching*



Matching.

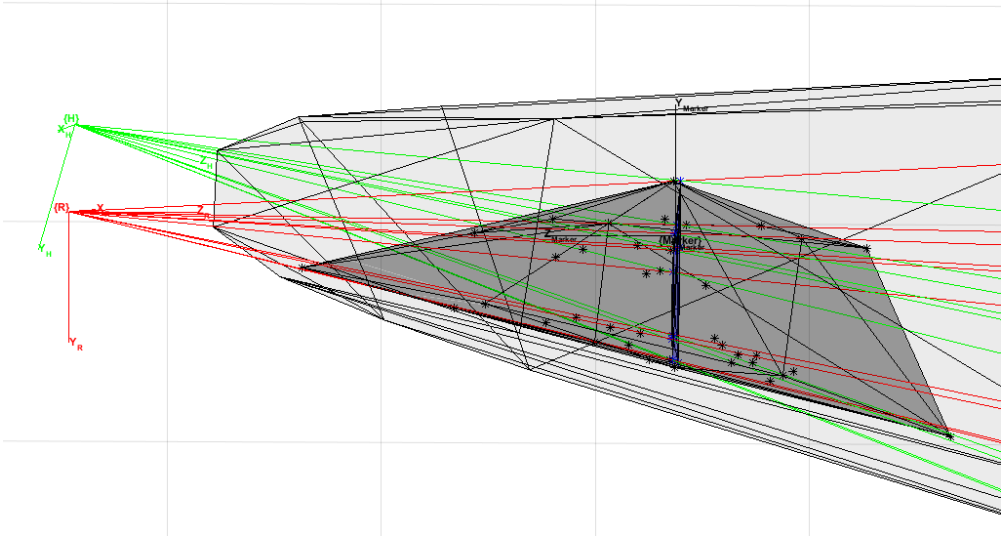
Results(II)

- *All the Simulated SLAM features some uncertainty was added in order to put the feature away of its real position.*



Representation of SLAM position estimation and corrected position estimation (View I). The gray shape is de Convex Hull of the simulated SLAM features, the blue one is the Convex Hull of the estimated Positions after Matching.

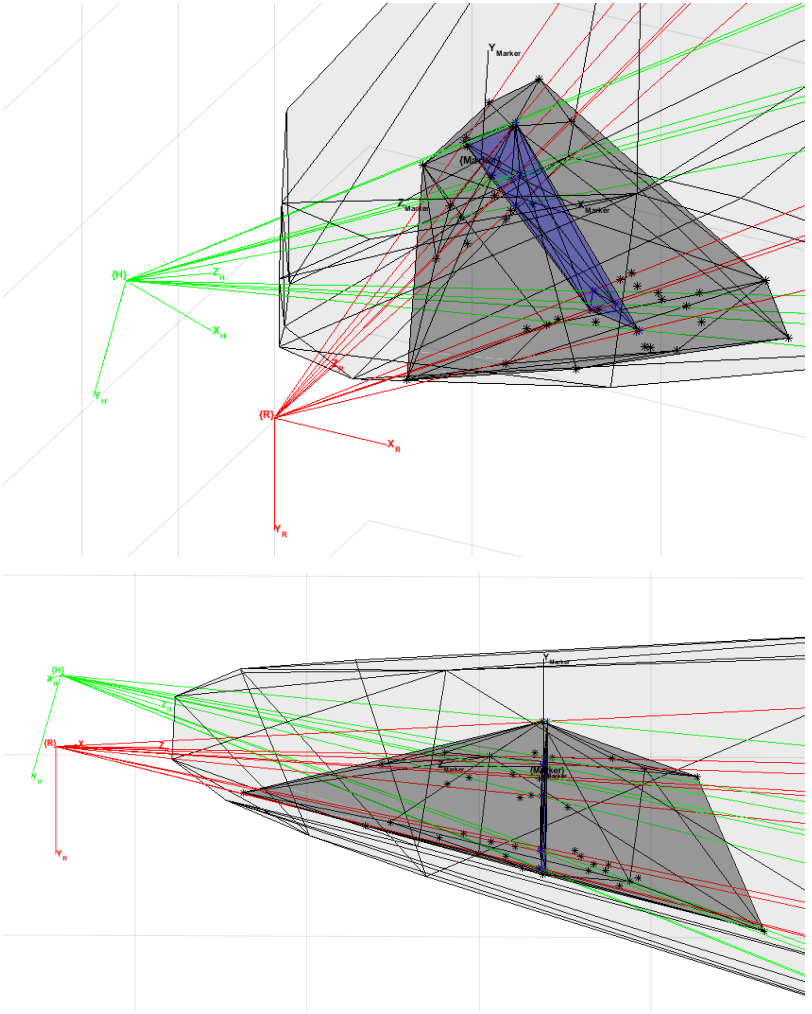
- *After estimate the feature position by matching, all features are in the same plane of the AR marker, good results for the correction.*



Representation of SLAM position estimation and corrected position estimation (View II). The gray shape is de Convex Hull of the simulated SLAM features, the blue one is the Convex Hull of the estimated Positions after Matching.

Conclusion and Further work

- *The geometry of the common field of view of two cameras is useful to threshold the space.*
- *Making use of it we reduce the number of features to look for, and the search space for each particular feature.*
- *Some work has to be done to link the approach with the Monocular SLAM algorithm.*
- *The performance of different descriptors should be tested for different type of sequences.*
- *A curve 3D region could be added to limit the volume where the features have more parallax than a given threshold, based in the baseline.*



Thanks!

Any question?