Detection of Critical Camera Configurations for Structure from Motion

Mario Michelini, Helmut Mayer
Institute of Applied Computer Science
Bundeswehr University Munich

Content

- Introduction
- Critical Camera Configurations
- Detection of Critical Camera Configurations
- Error Metrics
- Evaluation of Error Metrics
- Conclusions
- Future Work

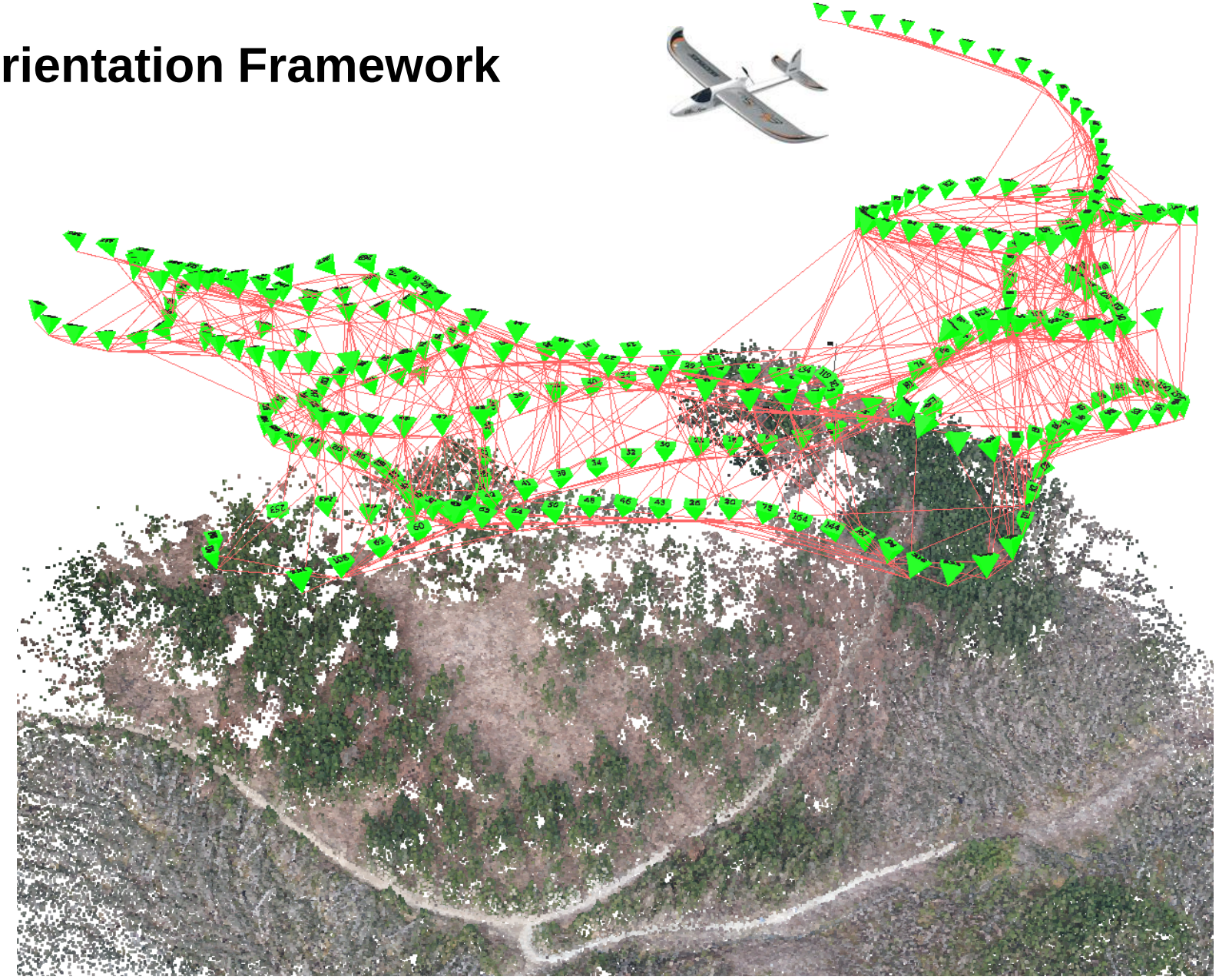
Introduction

Orientation Framework

- (Weakly) Calibrated Structure from Motion approach
- Can deal with unordered image sets and wide baselines between images.
- Builds on (Bartelsen et al., 2012, Mayer et al., 2012):
  - SIFT keypoints + cross-correlation + affine least squares matching
  - RANSAC + 5-point-algorithm + robust bundle adjustment
- Hierarchical merging of image sets into larger and larger sets

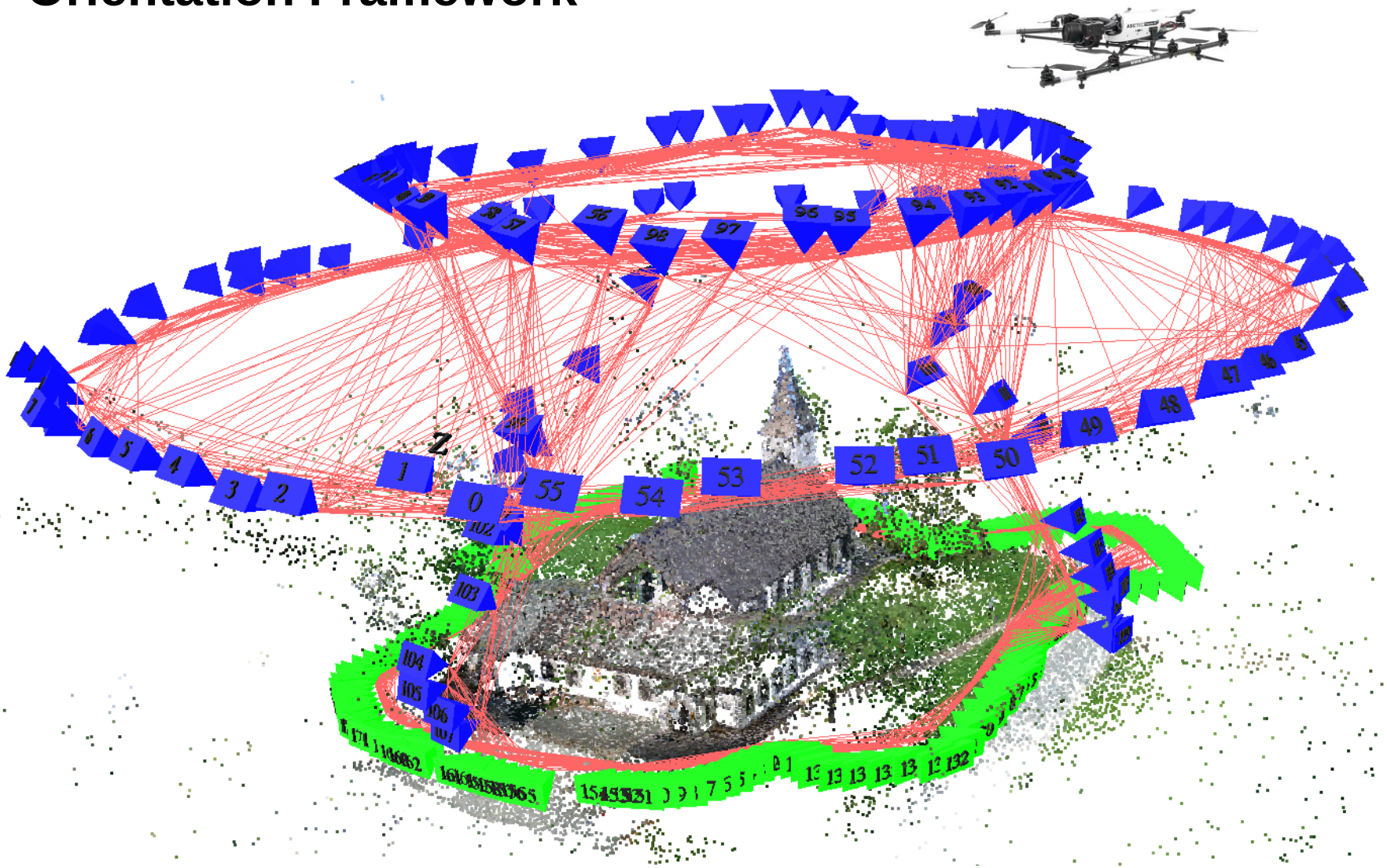
Introduction

Orientation Framework



Introduction

Orientation Framework



Introduction

Orientation Framework

- **Applications:**
  - Image interpretation (Nguaten et al., 2013)
  - Dense 3D reconstruction (Kuhn et al., 2013)

(Kuhn et. al., 2013)

Introduction

Orientation Framework

- **Applications:**
  - Image interpretation (Nguaten et al., 2013)
  - Dense 3D reconstruction (Kuhn et al., 2013)
- **Major limitation:** Not robust regarding critical camera configurations

Introduction

Orientation Framework

- **Applications:**
  - Image interpretation (Nguaten et al., 2013)
  - Dense 3D reconstruction (Kuhn et al., 2013)
- **Major limitation:** Not robust regarding critical camera configurations

➔ Improvement of the robustness concerning critical camera configurations by detecting them in advance

Critical Camera Configurations

- **Common assumptions** in most Structure from Motion approaches:
  - Scene is rigid.
  - Known scene structure
  - General camera configuration

- **Problems** can arise if:
  - Scene structure is unknown → *Structure degeneracy*
  - Arbitrary camera configuration → *Motion degeneracy*

Critical Camera Configurations

Structure Degeneracy

- Caused by planar scene structure
- Point correspondences related by homography
- Epipolar geometry is ambiguous.
- Geometry modeling based on epipolar geometry would lead to a random solution based on the inclusion of outliers.

$$F = [\mathbf{e'}]H \qquad \mathbf{e'} = \begin{pmatrix} 0 & 1 & e'_y \\ 1 & 0 & -e'_x \\ -e'_y & e'_x & 0 \end{pmatrix}$$

- Solution for
  - uncalibrated cameras: Model selection
  - calibrated cameras (our case): 5-point algorithm

Critical Camera Configurations

Motion Degeneracy

- No translation between views, i.e., pure rotational movement
- Point correspondences are related by the infinite homography:

$$H_\infty = K_2 R K_1^{-1}$$

- Epipolar geometry and 3D reconstruction are undefined.
- Analyzed in the context of keyframe selection approaches:
  - Selection of image pairs (*keyframes*) which are most suitable for reconstruction
  - Not designed for the detection of critical camera configurations

Critical Camera Configurations

- In triangulation-based Structure from Motion the **accuracy** is proportional to the ratio between the baseline length and the depth of the scene.

- **Critical camera configurations** are configurations where the baseline is:
  - Missing → Invalid reconstruction
  - Very short → Inaccurate reconstruction

➔ **Goal:** Detection of critical camera configurations which lead to poor or undefined intersection geometry

Detection of Critical Camera Configurations

- Most approaches (Pollefeys et al., 2002, Repko and Pollefeys, 2005, Beder and Steffen, 2006) were designed for keyframe selection and are, thus, less suitable for the detection of critical camera configurations.
- Formulation of the detection as a classification problem seems to be adequate.
- Error metrics are employed as classification features.
- Application of machine learning techniques to improve the classification

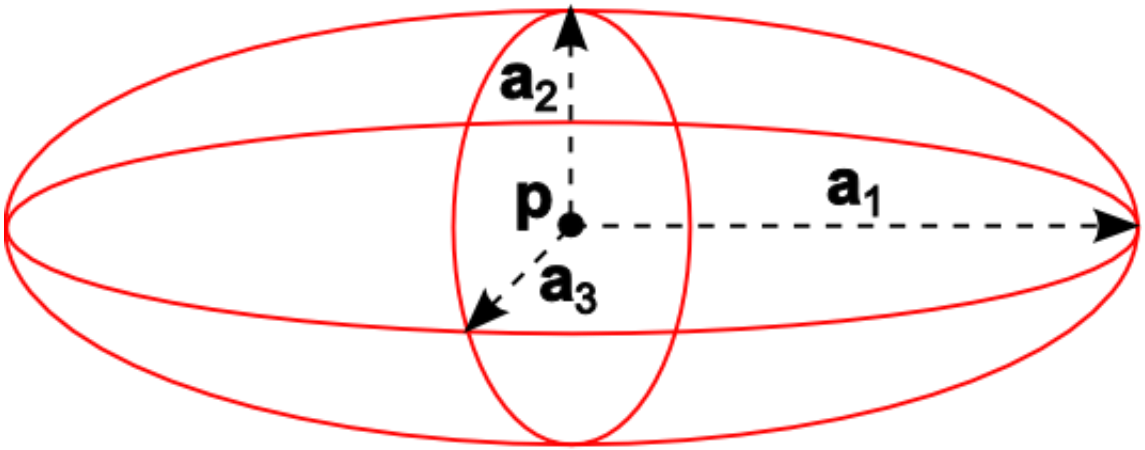
➔ Formulation of the detection of critical camera configurations as **binary classification problem** with **error metrics as classification features** and **AdaBoost** as classification algorithm

Error Metrics

Error Ellipsoid

- Covariance matrix C describes the uncertainty of a reconstructed 3D point **p**.
- Eigenvectors of C define the directions of the semi-axes $\mathbf{a}_i$ of the error ellipsoid.
- Eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ of C are the squares of the lengths of the semi-axes of the error ellipsoid:

$$\left|\mathbf{a}_i\right| = \sqrt{\lambda_i}$$



Error Metrics

Roundness

- Describes the roundness of the error ellipsoid.
- Defined as square root of the quotient of the smallest and the largest eigenvalue:

$$R = \sqrt{\frac{\lambda_3}{\lambda_1}} \qquad R = [0;1]$$

- Depends on the relative geometry of the two cameras and the feature positions.
- Mean roundness over all reconstructed points of an image pair is used for keyframe selection in (Beder and Steffen, 2006).
- Involves only two ellipsoid axes.

Error Metrics

Alternative Roundness

- Describes the roundness of the error ellipsoid.
- Defined as the quotient of the ellipsoid volume V and the volume V_K of the minimum circumscribed sphere:

$$K = \frac{V}{V_K} = \frac{\sqrt{\lambda_2 \lambda_3}}{\lambda_1} \qquad K = [0;1]$$

- Involves all ellipsoid axes.

Error Metrics

Alternative Roundness

- Describes the roundness of the error ellipsoid.
- Defined as the quotient of the ellipsoid volume V and the volume V_K of the minimum circumscribed sphere:

$$K = \frac{V}{V_K} = \frac{\sqrt{\lambda_2 \lambda_3}}{\lambda_1} \qquad K = [0;1]$$

- Involves all ellipsoid axes.

Volume

- Describes the volume of the error ellipsoid:

$$V = \frac{4}{3}\pi\sqrt{\lambda_1 \lambda_2 \lambda_3} = \frac{4}{3}\pi\sqrt{\det(C)} \qquad V \in \mathbb{R}^+$$

- Can be computed directly from the covariance matrix C without eigendecomposition.

Error Metrics

Sphericity

- Measures how spherical the error ellipsoid is.
- Defined as the ratio of the surface area of a sphere with the same volume as the ellipsoid to the surface area of the ellipsoid:

$$S = \frac{\pi^{\frac{1}{3}}(6V)^{\frac{2}{3}}}{O} = \frac{4\pi\sqrt{\lambda_1\lambda_2\lambda_3}}{O} \qquad S = [0;1]$$

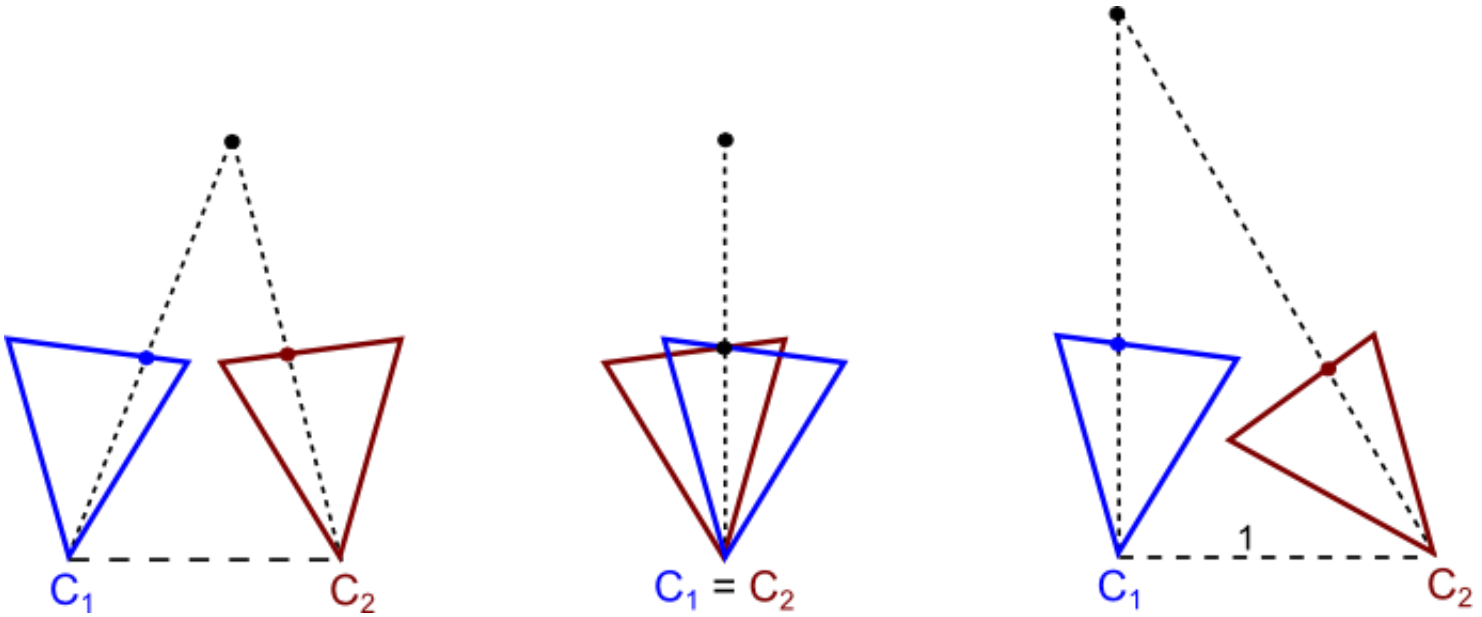
$$O \approx 4\pi\left(\frac{\sqrt{\lambda_1\lambda_2}^{p} + \sqrt{\lambda_1\lambda_3}^{p} + \sqrt{\lambda_2\lambda_3}^{p}}{3}\right)^{1/p} \qquad p = 1.6$$

- O is an approximation of the ellipsoid surface area.
- p = 1.6 leads to a maximum relative error of 1.178%.

Error Metrics

Depth

- Measures the relative depth of the reconstructed 3D point.
- Proportional to the baseline
- Independent of the covariance matrix
- Depth of the reconstructed 3D points for critical and general camera configurations often differ significantly.



general degenerated pseudo

Error Metrics

Summary

- Metrics based on the shape of error ellipsoid:
  - Roundness R
  - Volume V
  - Alternative Roundness K
  - Sphericity S
- Shape independent metric: Depth D of the reconstructed 3D point

All metrics are computed for each reconstructed 3D point and the median is used as global metric.

Mean of the roundness R_{mean} as proposed in (Beder and Steffen, 2006) is employed for comparison.

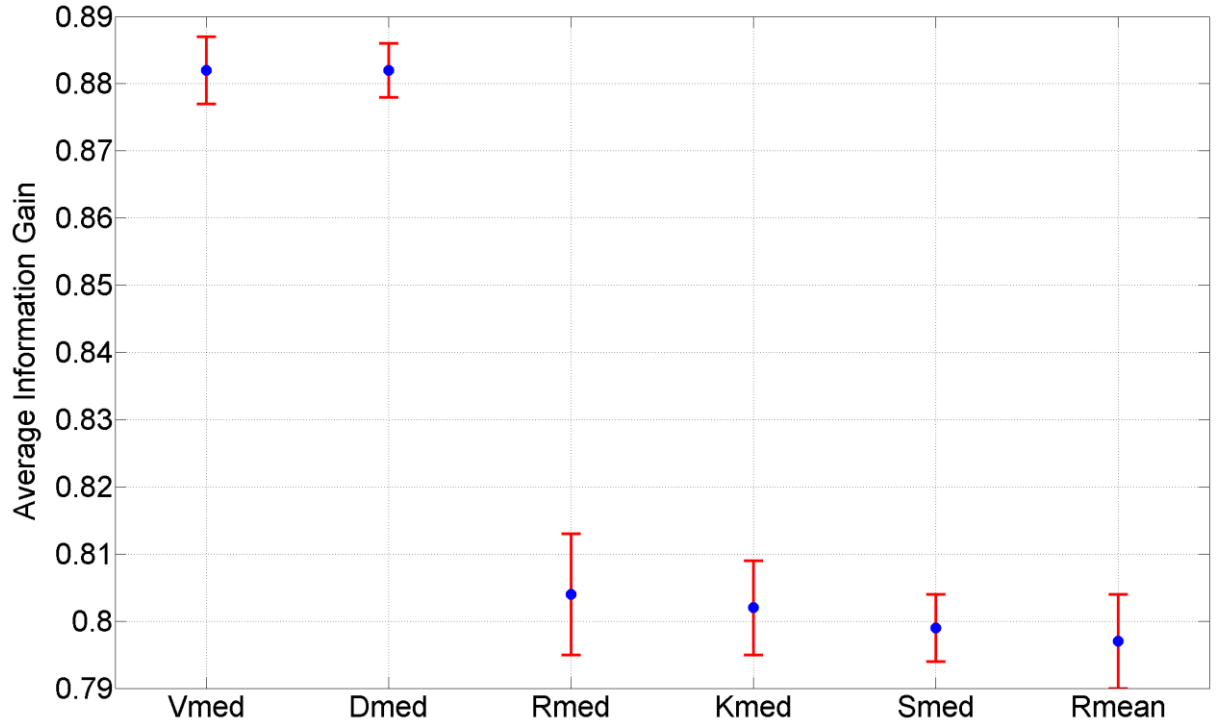
Evaluation of Error Metrics

- Training data:
  - 1500 image pairs acquired by handheld cameras and cameras on small Unmanned Aerial System (UAS)
  - 70% positive samples: Pairs with general camera configuration
  - 30% negative samples: Pairs with motion degeneracy
- Classifier's performance is measured using precision and recall combined in F-score.
- Evaluation score is computed from mean of stratified cross validation with 10 folds repeated 10 times.

Evaluation of Error Metrics

Feature Comparison

- Suitability as a single classification feature measured using information gain:



- V_{med} and D_{med} perform best with small variations.
- Roundness-based features behave all similar.

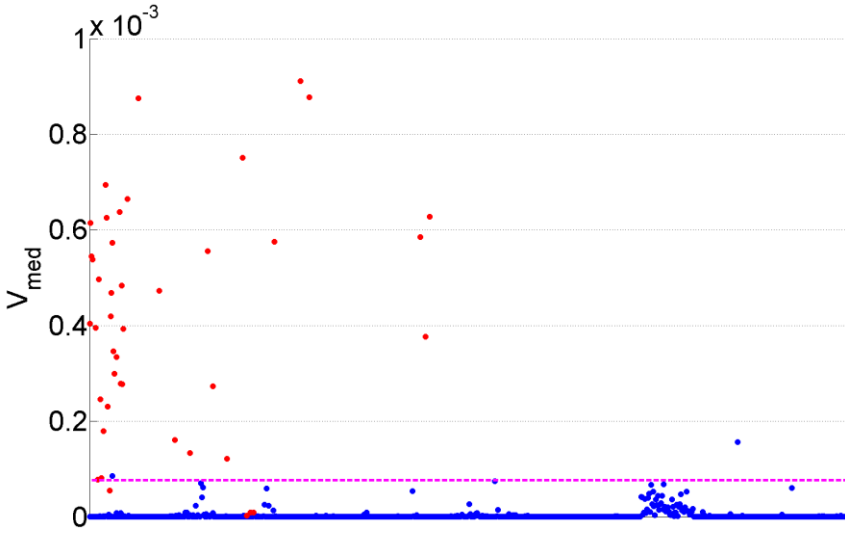
Evaluation of Error Metrics

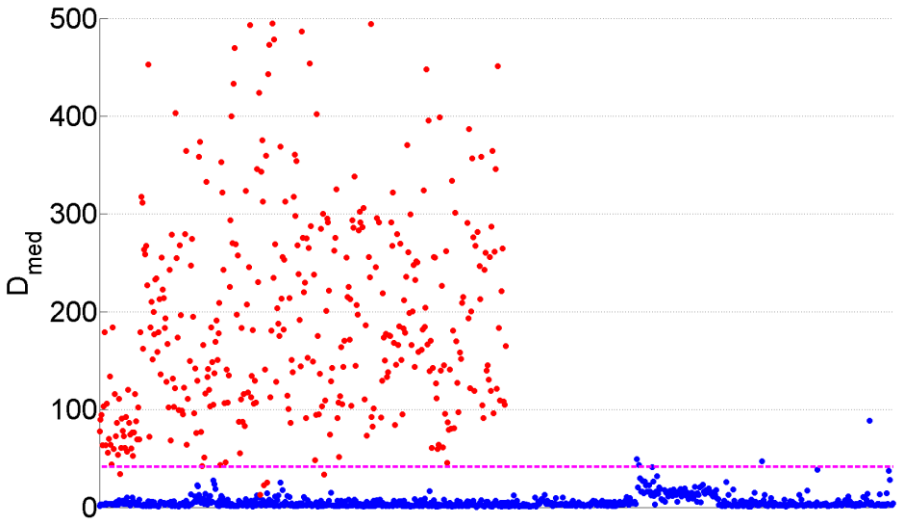
Feature Comparison

- Suitability of a single feature for classification
- Evaluation based on simple decision stump (i.e. threshold selection):

Feature	V_{med}	D_{med}	K_{med}	R_{med}	R_{mean}	S_{med}
F-score	0.9940	0.9915	0.9777	0.9792	0.9738	0.9787

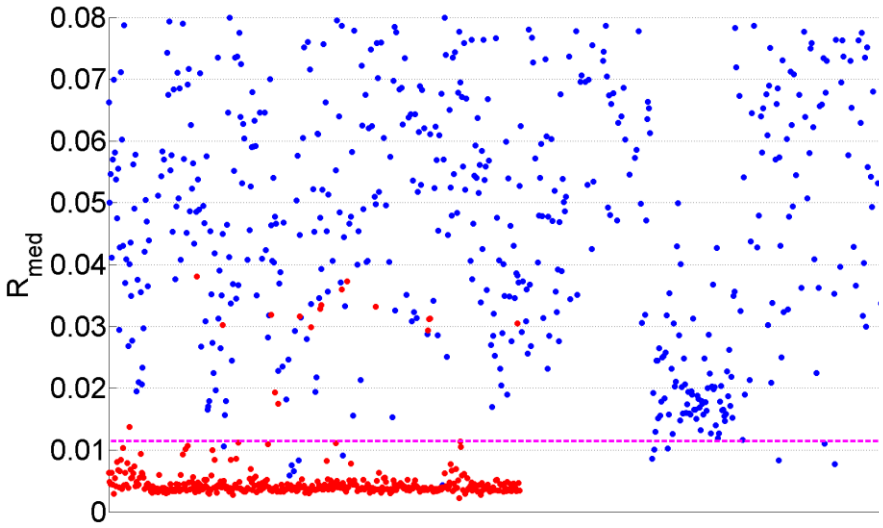
- $(V_{med}, D_{med}) \rightarrow (K_{med}, R_{med}, S_{med}) \rightarrow R_{mean}$

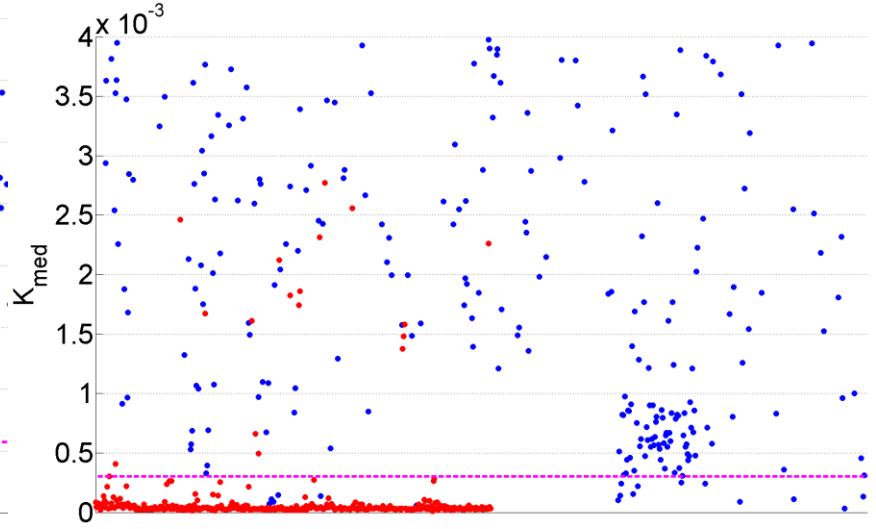


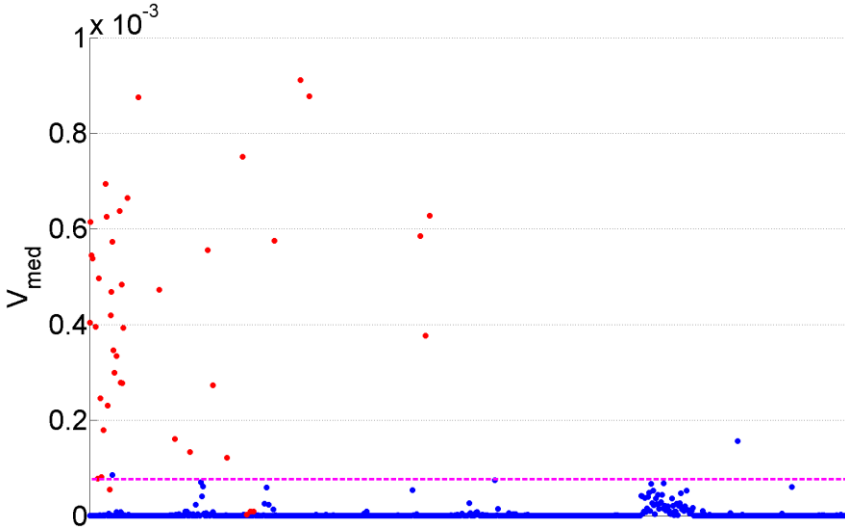


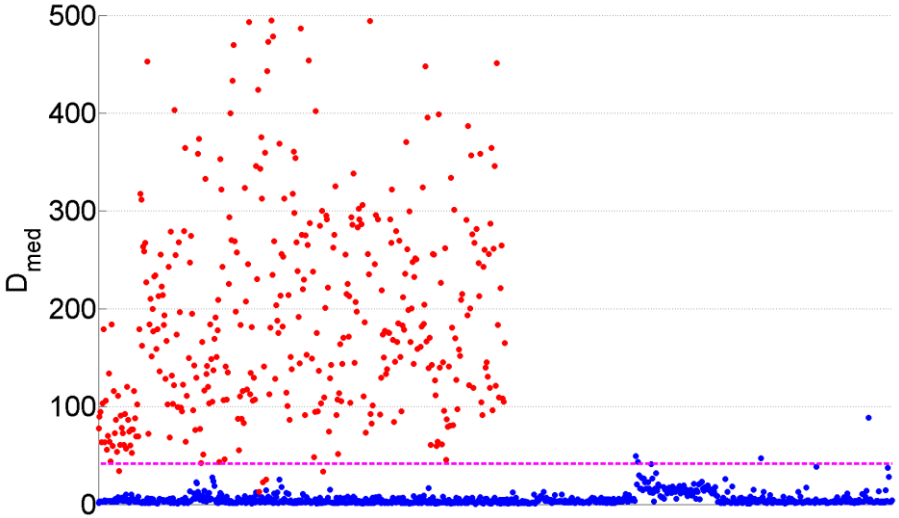
Evaluation of Error Metrics

Feature Comparison







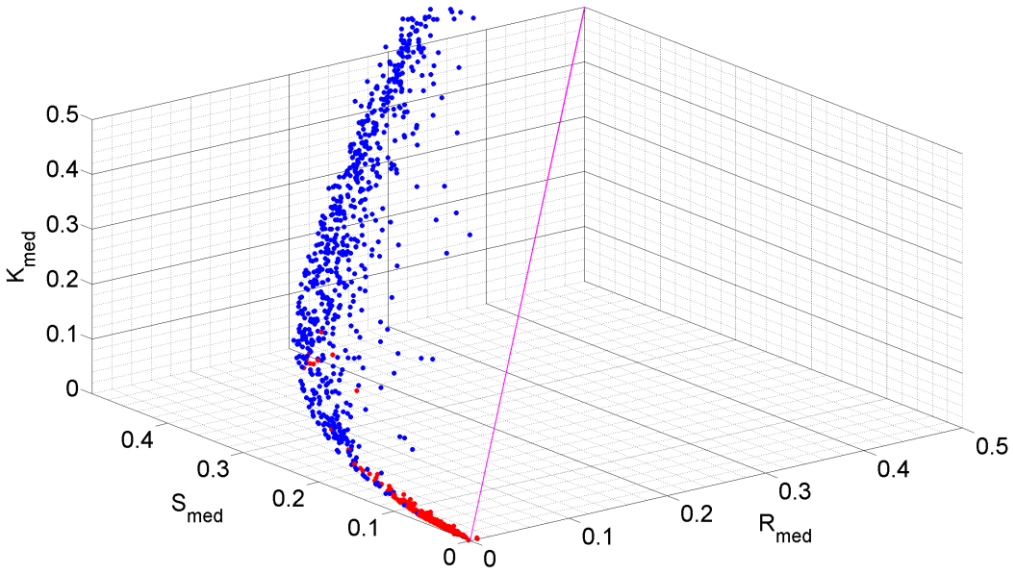


Evaluation of Error Metrics

Feature Comparison

- Roundness-based features seem to be similar.
- Comparison of roundness-based features using correlation:

Features	Pearson	Spearman
R_{med} S_{med}	0.95	0.99
R_{med} K_{med}	0.84	0.96
S_{med} K_{med}	0.77	0.95



- Features are correlated pairwise, not linearly but rather quadratically.
- There exist also a non-linear correlation between all three features.

Evaluation of Error Metrics

Feature Comparison – Results

- Features V_{med} and D_{med} show best behavior concerning suitability as single classification features.
- Roundness-based features are correlated and thus it is sufficient to use only one of them:
  - S_{med} is only an approximation and costly to compute.
  - K_{med} and R_{med} are simple, but K_{med} involves the whole ellipsoid shape.
- Mean roundness R_{mean} performs worst and, thus, median roundness R_{med} should be used instead.

Evaluation of Error Metrics

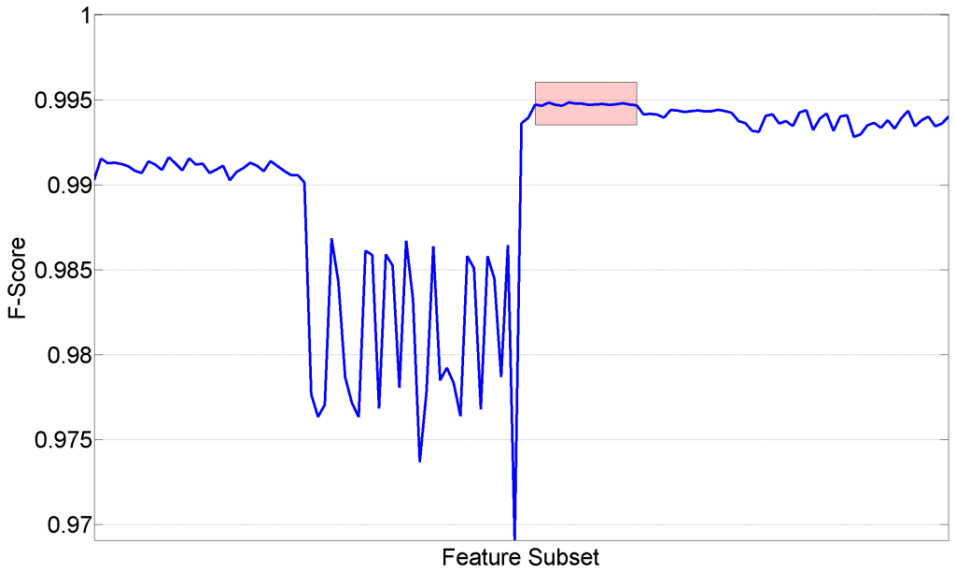
Feature Selection

- Feature comparison has considered only the performance of a single feature.
- Feature selection chooses the best feature subset for a specific classifier.
- Feature selection performed by exhaustive search with
  - all non empty sets of the feature set's power set
  - F-score as evaluation measure
  - AdaBoost as classifier with different number of iterations

Evaluation of Error Metrics

Feature Selection – Results

- Performance of the feature subsets is not very sensitive to the number of boost iterations.
- Combination of V_{med} and D_{med} gives the best result.
- Further improvement by involving one of the roundness-based features.
- Good performance comes primarily from V_{med}.



Feature Subset	F-score
V_{med} D_{med} K_{med}	0.9947
V_{med} D_{med} R_{med}	0.9944
V_{med} D_{med} S_{med}	0.9943
V_{med} D_{med}	0.9939
V_{med}	0.9936
D_{med}	0.9903

Conclusions

- Orientation framework for unordered image sets has been presented.
- Problems in Structure from Motion have been discussed.
- Detection of critical camera configurations has been formulated as binary classification problem.
- Various error metrics have been presented and analyzed concerning their suitability as classification features for AdaBoost.
- Three-dimensional feature vector (V_{med},D_{med},K_{med}) leads to the best classification error of less than 1%.

Future Work

- Integration of the detection of critical camera configuration into orientation framework to improve the handling of unordered image sets.
- Evaluation of other classification algorithms
- Formulation and evaluation of detection of critical camera configurations as multiclass classification problem

References

Bartelsen, J., Mayer, H., Hirschmüller, H., Kuhn, A. and Michelini, M., 2012. *Orientation and Dense Reconstruction from Unordered Wide Baseline Image Sets*. Photogrammetrie – Fernerkundung – Geoinformation 4/12, pp. 421–432.

Bader, C. and Steffen, R., 2006. *Determining an Initial Image Pair for Fixing the Scale of a 3D Reconstruction from an Image Sequence*. In: Pattern Recognition – DAGM 2006, Springer-Verlag, Berlin, Germany, pp. 657-666

Kuhn, A., Hirschmüller, H. and Mayer, H., 2013. *Multi-Resolution Range Data Fusion for Multi-View Stereo Reconstruction*. In: Pattern Recognition, Lecture Notes in Computer Science, Vol. 8142, Springer Berlin Heidelberg, pp. 41–50.

Nguaten, W., Drauschke, M., Mayer, H., 2013. *Roof Reconstruction from Point Clouds using Importance Sampling*. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Science II (3/W3), pp. 79–78.

Mayer, H., Bartelsen, J., Hirschmüller, H. and Kuhn, A., 2012. *Dense 3D Reconstruction from Wide Baseline Image Sets*. In: Real-World Scene Analysis 2011, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Germany, pp. 285–304.

Pollefeys, M., Verbiest, F. and Gool, L. J. V., 2002b. *Surviving Dominant Planes in Uncalibrated Structure and Motion Recovery*. In: 7th European Conference on Computer Vision – Part II, Springer-Verlag, pp. 837–851.

Repko, J. and Pollefeys, M., 2005. 3D Models from Extended Uncalibrated Video Sequences: *Addressing Key-Frame Selection and Projective Drift*. In: 5th International Conference on 3-D Digital Imaging and Modeling, IEEE Computer Society, pp. 150–157.